

RAFAELA BOEIRA GIORDANO

DO JORNAL À CIÊNCIA:
a Hemeroteca Digital Brasileira como
fonte de informação para a pesquisa científica

Tese de Doutorado
Junho de 2016



**UNIVERSIDADE FEDERAL DO RIO DE JANEIRO – UFRJ
ESCOLA DE COMUNICAÇÃO – ECO
INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA
E TECNOLOGIA – IBICT
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO – PPGCI**

RAFAELA BOEIRA GIORDANO

**DO JORNAL À CIÊNCIA:
a Hemeroteca Digital Brasileira como
fonte de informação para a pesquisa científica**

RIO DE JANEIRO

2016

Rafaela Boeira Giordano

**DO JORNAL À CIÊNCIA:
a Hemeroteca Digital Brasileira como
fonte de informação para a pesquisa científica**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Informação, convênio entre o Instituto Brasileiro de Informação em Ciência e Tecnologia e a Universidade Federal do Rio de Janeiro/Escola de Comunicação, como requisito parcial para a obtenção do título de Doutora em Ciência da Informação.

Orientador: Jorge Calmon de Almeida Biolchini

Rio de Janeiro

2016

G494d Giordano, Rafaela Boeira.

Do jornal à ciência: a hemeroteca digital brasileira como fonte de informação para a pesquisa científica / Rafaela Boeira Giordano. – Rio de Janeiro: 2016.

239 f. : il., 30cm.

Tese (Doutorado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Instituto Brasileiro de Informação em Ciência e Tecnologia, Universidade Federal do Rio de Janeiro, Escola de Comunicação, 2016.

Orientador: Jorge Calmon de Almeida Biolchini.

1. Hemeroteca Digital. 2. Biblioteca Nacional. 3. Recuperação da informação. 4. Jornalismo. 5. Ciência da Informação – Tese. I. Biolchini, Jorge Calmon de Almeida (Orient.). II. Universidade Federal do Rio de Janeiro, Escola de Comunicação. III. Instituto Brasileiro de Informação em Ciência e Tecnologia, Programa de Pós-Graduação em Ciência da Informação. IV. Título

CDD 020

Rafaela Boeira Giordano

**DO JORNAL À CIÊNCIA:
a Hemeroteca Digital Brasileira como
fonte de informação para a pesquisa científica**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Informação, convênio entre o Instituto Brasileiro de Informação em Ciência e Tecnologia e a Universidade Federal do Rio de Janeiro/Escola de Comunicação, como requisito parcial à obtenção do título de Doutora em Ciência da Informação.

Aprovada em: 20 de junho de 2016

Prof. Dr. Jorge Calmon de Almeida Biolchini (orientador)
PPGCI/IBICT-ECO/UFRJ

Prof^a. Dr^a. Gilda Olinto de Oliveira
PPGCI/IBICT-ECO/UFRJ

Prof^a. Dr^a. Rosali Fernandez de Souza
PPGCI/IBICT-ECO/UFRJ

Prof^a. Dr^a. Joice Cardoso Ennes de Souza
Universidade Federal Fluminense (UFF) / Departamento de Ciência da Informação

Prof. Dr. Marcos Luiz Cavalcanti de Miranda
Universidade Federal do Estado do Rio de Janeiro (UNIRIO) / Programa de Pós-Graduação em Biblioteconomia

AGRADECIMENTOS

Muito obrigada:

Ao professor Jorge Biolchini pela orientação.

A Claudio Werneck, por estar ao meu lado durante os anos de doutorado.

A minha querida amiga Angela Monteiro Bettencourt pela sugestão do tema.

À querida professora Lena Vania Ribeiro Pinheiro por nos contagiar com seu amor pela Ciência da Informação. À professora Gilda Olinto por ser tão sincera nas avaliações e interessada no desenvolvimento dos alunos. Aos professores Gustavo Saldanha e Marco Schneider por nos orientarem a desconfiar do óbvio. À Janete Desidério por ser sempre tão atenciosa e solícita. A Sebastião Neves da Silva, o Tião, pelos sorrisos.

A Regina Lourenço e Luana Sales pelos modelos de projetos. À bibliotecária Sarah Miglioli pela ajuda na recuperação de textos. A Antonio Victor Botão, Ana Rosa Ribeiro, Neusa Cardim e Tatiana Almeida pelo melhor CI. A Eliane Costa, Gabriela Portolan, Juliana Zanotto, Lisély Luz, Márcia Feijó, Marianna Zattar e Marie Mompezat pelo apoio e amizade durante esse longo processo. A Julia Wiśniewska e Ursula Dannemann *für die ständige Unterstützung*. A Daise Ribeiro Carpes, Monica Heitz, Rita Meyer, Rodrigo Batagello e Silvia Lopes pelas dicas e incentivo. Ao seu José Gerpes Paganella e à dona Marli Konig dos Santos pelas melhores memórias da infância em meio a livros e gibis. À Maria Luiza pelas pistas. A Luci Giordano e Regina Helena Werneck pelo pensamento positivo.

A Milena Ambrósio Telles e Emília Barroso Cruz por uma série de motivos.

Ao professor Robert Wegner, do Programa de Pós-Graduação em História das Ciências e da Saúde (PPGHCS) da Fiocruz, pelo interesse e disposição em ajudar na busca de voluntários para a pesquisa. Aos alunos e egressos do PPGHCS/Fiocruz pela colaboração.

A Christa Müller (Biblioteca Nacional da Áustria), Deborah Thomas (Library of Congress), Emerson Vandy (Biblioteca Nacional da Nova Zelândia), Laura Rodriguez (Biblioteca Nacional da Costa Rica), Lola Rodríguez (Biblioteca Nacional da Espanha), Meinrad Büchel (Biblioteca Nacional de Liechtenstein), Marc D'Hoore, Veronique Wese e Maud Henry (Biblioteca Real da Bélgica) pelas informações sobre as hemerotecas.

Aos atenciosos professores da Universidade de Caxias do Sul (UCS) João Paulo Borges da Silveira, Marcos Leandro Freitas Hübner e Michele Marques Baptista.

Aos professores Gilda Olinto, Joice Cardoso Ennes de Souza, Marcos Luiz Cavalcanti de Miranda, Naira Christofolletti Silveira, Ricardo Medeiros Pimenta e Rosali Fernandez de Souza por aceitarem fazer parte da minha trajetória acadêmica.

RESUMO

GIORDANO, Rafaela Boeira. *Do jornal à ciência: a hemeroteca digital brasileira como fonte de informação para a pesquisa científica*. 2016. 239 f. Tese (Doutorado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Instituto Brasileiro de Informação em Ciência e Tecnologia, Universidade Federal do Rio de Janeiro, Escola de Comunicação, PPGCI, Rio de Janeiro, 2016.

Hemerotecas digitais foram criadas pelas Bibliotecas Nacionais nos últimos 15 anos a partir da digitalização de jornais e revistas antigos. A iniciativa propicia a preservação dos acervos e o acesso facilitado à coleção, via “Web”, a qualquer hora e de qualquer lugar. A informação, objeto de estudo da Ciência da Informação e matéria-prima do Jornalismo, fica mais acessível à medida que mais e mais publicações periódicas são digitalizadas por diferentes Bibliotecas Nacionais. Neste estudo, buscou-se analisar as características de hemerotecas digitais criadas por Bibliotecas Nacionais, apresentando assim o estado da arte do desenvolvimento dessas coleções. A partir das questões de pós-graduandos da área de História da Saúde foram realizadas buscas na Hemeroteca Digital Brasileira da Biblioteca Nacional do Brasil, que permitiram avaliar a hemeroteca como fonte de informação científica. A oferta *on-line* de coleções de jornais e revistas pelas Bibliotecas Nacionais facilitou e solidificou o uso de notícias como matéria-prima em investigações científicas, seja como elemento-base da pesquisa, como um complemento ou para gerar *insights*. A partir do exame de treze hemerotecas digitais, observou-se ainda que cada uma adota critérios próprios de seleção do material a ser digitalizado e de difusão dessa informação, mas a representação tem como base tradicionais fundamentos da Organização do Conhecimento. A grande inovação tem sido a adoção do uso de reconhecimento óptico de caracteres (OCR), que permite aos usuários buscar e recuperar informações nos conteúdos das publicações.

Palavras-chave: Hemeroteca Digital. Biblioteca Nacional. Busca de Informação. Jornalismo. Pesquisa Histórica. História da Saúde.

ABSTRACT

GIORDANO, Rafaela Boeira. *Do jornal à ciência: a hemeroteca digital brasileira como fonte de informação para a pesquisa científica*. 2016. 239 f. Tese (Doutorado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Instituto Brasileiro de Informação em Ciência e Tecnologia, Universidade Federal do Rio de Janeiro, Escola de Comunicação, PPGCI, Rio de Janeiro, 2016.

Digital newspaper libraries have been created by National Libraries for the past 15 years from the digitization of old newspapers and magazines. The initiative promotes the preservation of collections and the access through the “Web” anytime from anywhere. Information, the object of study of Information Science and the main element of Journalism, becomes more accessible as more and more journals are scanned by different National Libraries. This study aimed to analyze the features of digital newspaper libraries created by National Libraries, presenting the state of the art of these collections. From the research themes of graduate students of Health History searches were carried out in the Brazilian Digital Newspaper Library of Brazilian National Library, which made possible to evaluate the newspaper library as a source of scientific information. The offer of online newspapers and magazines by National Libraries facilitated the use of news as an important element in scientific research, or as the base of a research, or still as an insight. From the examination of thirteen digital national newspaper libraries, it was observed that each one adopts its own criteria for selection of material to be scanned and to disseminate, but the information representation has traditional foundations in Knowledge Organization. The major innovation detected was the adoption of optical character recognition (OCR), which allows users to search and retrieve information in the publications content.

Keywords: Digital Newspaper Library. National Library. Information Search. Journalism. Historical Research. History of Health.

LISTA DE FIGURAS

Figura 1 –	Percurso da revisão bibliográfica	25
Figura 2 –	OC/RC, OI/RI	33
Figura 3 –	Desenho de Otlet mostra interação entre documentação e telecomunicação	40
Figura 4 –	Relações hierárquicas em um tesouro	44
Figura 5 –	Processos de trabalho dos SRI com a informação	76
Figura 6 –	Processo de busca e recuperação da informação	79
Figura 7 –	Modelo <i>Berry-picking</i>	80
Figura 8 –	Representação do resultado de expressão booleana conjuntiva (and)	81
Figura 9 –	Representação do resultado de uma busca booleana disjuntiva (or)	82
Figura 10 –	Resultado de busca negativa (not)	82
Figura 11 –	Resultado de busca booleana com o operador not	82
Figura 12 –	Página inicial da Trapalanda, Biblioteca Digital Argentina	106
Figura 13 –	Página de busca da Trapalanda	107
Figura 14 –	Opção de listar o acervo por autor, data, local e assunto	108
Figura 15 –	Página de resultados de busca na seção Periódicos da Trapalanda	108
Figura 16 –	Página recuperada na Trapalanda	109
Figura 17 –	Página inicial do ANNO	111
Figura 18 –	Busca básica do ANNO	112
Figura 19 –	Busca avançada do ANNO	112
Figura 20 –	Resultado da busca apresentado em forma de lista	113
Figura 21 –	Página em português, de <i>O Spectador Brasileiro</i> , recuperada no ANNO	114
Figura 22 –	Mapa interativo do ANNO com os locais em que foram encontrados os jornais	115
Figura 23 –	Página inicial da seção <i>Digitised newspapers and more</i> do Trove	118
Figura 24 –	Página de resultados do Trove	119
Figura 25 –	Detalhe da apresentação dos resultados no Trove	120
Figura 26 –	Resultado de busca empreendida no Trove	121
Figura 27 –	Página inicial da coleção de jornais da KBR	124
Figura 28 –	Página com resultados da KBR em forma de lista	125
Figura 29 –	Página com um dos resultados da KBR selecionados na lista geral	125

Figura 30 –	Página de busca nos periódicos costa-riquenhos	128
Figura 31 –	Lista de resultados obtidos no SINABI	128
Figura 32 –	Página de busca da Hemeroteca Digital da Bib. Nacional da Espanha ...	131
Figura 33 –	Página de resultados da hemeroteca espanhola	132
Figura 34 –	Página recuperada na Hemeroteca Digital espanhola	133
Figura 35 –	Página inicial do <i>Chronicling America</i>	135
Figura 36 –	Página recuperada no <i>Chronicling America</i>	136
Figura 37 –	Opções de visualização do material recuperado no CA	136
Figura 38 –	Página inicial da Gallica, a Biblioteca Digital da BN francesa	137
Figura 39 –	Página inicial da seção <i>Presse et Revues</i> da Gallica	138
Figura 40 –	Busca simples para qualquer seção do Gallica	139
Figura 41 –	Resultado da busca realizada por título na Gallica	140
Figura 42 –	Visualização de uma página recuperada no <i>app</i> da Gallica	141
Figura 43 –	Página de busca da eLiechtensteinesia	142
Figura 44 –	Página de busca básica da Hemeroteca Digital do México – por ano	145
Figura 45 –	Página de busca básica Hemeroteca Digital do México – por título	146
Figura 46 –	Página de busca básica Hemeroteca Digital do México – por localização	147
Figura 47 –	Página inicial do <i>Papers Past</i>	149
Figura 48 –	Resultados da busca no <i>Papers Past</i>	149
Figura 49 –	Visualização do resultados da busca no <i>Papers Past</i>	150
Figura 50 –	Página inicial do <i>The British Newspaper Archive</i>	151
Figura 51 –	<i>National Newspaper Building</i> do BNA, em Boston Spa	153
Figura 52 –	Resultados da busca no BNA	154
Figura 53 –	Período de cobertura das hemerotecas	156
Figura 54 –	Leiaute das páginas de busca das hemerotecas estudadas	159-161
Figura 55 –	Detalhe da apresentação de resultados no ANNO	162
Figura 56 –	Detalhe da apresentação de resultados no Trove	163
Figura 57 –	Detalhe da apresentação de resultados no BelgicaPress	163
Figura 58 –	Detalhe da apresentação de resultados na hemeroteca espanhola	164
Figura 59 –	Detalhe da apresentação de resultados no <i>Chronicling America</i>	164
Figura 60 –	Detalhe da apresentação de resultados na Gallica	165
Figura 61 –	Detalhe da apresentação de resultados na e-Lichtensteinesia	165

Figura 62 – Detalhe da apresentação de resultados na hemeroteca mexicana	166
Figura 63 – Detalhe da apresentação de resultados no <i>Paper Past</i>	166
Figura 64 – Detalhe da apresentação de resultados no BNA	167
Figura 65 – Primeira página do <i>Correio Braziliense</i>	170
Figura 66 – Página inicial da Hemeroteca Digital Brasileira	192
Figura 67 – Busca avançada da Hemeroteca Digital Brasileira	193
Figura 68 – Apresentação dos resultados na Hemeroteca Digital Brasileira	218
Figura 69 – Modelo de conhecimento para reflexão sobre a busca de informação em jornais e revistas	221

LISTA DE QUADROS

Quadro 1 –	Comparação entre linguagens de indexação e vocabulário controlado ...	54
Quadro 2 –	Classificações dos textos que compõem um jornal	65
Quadro 3 –	Proposta de classificação de Marques Beltrão	66
Quadro 4 –	Gerações dos SRI	77
Quadro 5 –	Táticas de busca da informação	84-86
Quadro 6 –	Desempenho do OCR, fatores e indicações	98-99
Quadro 7 –	Bibliotecas Nacionais com hemerotecas digitais	105
Quadro 8 –	Parâmetros/metadados usados nas buscas simples e avançada	158
Quadro 9 –	Formato da apresentação dos resultados	162
Quadro 10 –	Recursos usados na filtragem dos resultados	167
Quadro 11 –	Compartilhamento de resultados via “Web”	168
Quadro 12 –	Perguntas básicas do Jornalismo alinhadas às categorias de Ranganathan	221
Quadro 13 –	Aplicação do modelo de conhecimento às buscas empreendidas na HDB	223
Quadro 14 –	Resumo das buscas empreendidas na HDB	224-225

LISTA DE TABELAS

Tabela 1 –	Data de lançamento das hemerotecas digitais	155
Tabela 2 –	Número de títulos e de páginas digitalizados	157

LISTA DE SIGLAS

AACR	Anglo-American Cataloguing Rules
ABNT	Associação Brasileira de Normas Técnicas
ALA	American Library Association
ALTO	Analyzed Layout and Text Object
ANDP	Australian Newspaper Digitisation Program
ANJ	Associação Nacional de Jornais
BDTD	Biblioteca Digital de Teses e Dissertações
BN	Biblioteca Nacional
BNA	British Newspaper Archive
BNDigital	Biblioteca Nacional Digital
BNE	Biblioteca Nacional de España
BNF	Bibliothèque Nationale de France
CA	Chronicling America
CALCO	Catálogo Legível por Computador
CCAA	Código de Catalogação Anglo-Americano
CDC	Curso de Especialização em Informação e Documentação
CDD	Classificação Decimal de Dewey
CDU	Classificação Decimal Universal
CEGES	Centro de Estudos Históricos sobre Guerra e Sociedade Contemporânea da Bélgica
CMS	Content Management System
COC	Casa de Oswaldo Cruz
Conarq	Conselho Nacional de Arquivos
DIP	Departamento de Imprensa e Propaganda
DLF	Digital Library Federation
EAC	Encoded Archival Context
ELSP	Escola Livre de Sociologia e Política
FID	Federação Internacional de Informação e Documentação
Fiocruz	Fundação Oswaldo Cruz
FNB	Fundação Biblioteca Nacional
FRBR	Functional Requirements for Bibliographic Records
HDB	Hemeroteca Digital Brasileira

IBBD	Instituto Brasileiro de Bibliografia e Documentação
IBICT	Instituto Brasileiro de Informação em Ciência e Tecnologia
IBTE	Instituto Behring de Terapêutica Experimental
IFLA	International Federation of Library Associations and Institutions
IIB	Institut International de Bibliographie
IPHAN	Instituto do Patrimônio Histórico e Artístico Nacional
ISBD	Descrição Bibliográfica Internacional Normalizada
ISSN	International Standard Serial Number
KRB	Koninklijke Bibliotheek van België (Biblioteca Real da Bélgica)
LC	Library of Congress (Biblioteca do Congresso dos Estados Unidos da América)
MARC	Machine Readable Cataloging
METS	Metadata Encoding and Transmission Standard
NDNP	National Digital Newspaper Program
MIX	Metadata for Images in XML Standard
MODS	Metadata Object Description Schema
NEH	National Endowment for the Humanities
NISO	National Information Standards Organization
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
OC	Organização do Conhecimento
OCR	Optical Character Recognition
OI	Organização da Informação
ONU	Organização das Nações Unidas
OPAC	Online Public Access Catalog
PDF	Portable Document Format
PLANO	Plano Nacional de Microfilmagem de Periódicos Brasileiros
PMEST	Personalidade, Matéria, Energia, Espaço e Tempo
PPGCI	Programa de Pós-Graduação em Ciência da Informação
PPGHCS	Programa de Pós-Graduação em História das Ciências e da Saúde
PUC	Pontifícia Universidade Católica
RC	Representação do Conhecimento
RDA	Resource Description and Access
RI	Representação da Informação
SIC	Serviço de Intercâmbio de Catalogação

Sinabi	El Sistema Nacional de Bibliotecas da Costa Rica
SNDM	Serviço Nacional de Doenças Mentais
SOC	Sistemas de Organização do Conhecimento
SRI	Sistema de Recuperação da Informação
SUS	Sistema Único de Saúde
TI	Tecnologia de Informação
TIFF	Tagged Image File Format
UCLA	University of California, Los Angeles
UFF	Universidade Federal Fluminense
UFRJ	Universidade Federal do Rio de Janeiro
UFSC	Universidade Federal de Santa Catarina
UN	The United Nations
UNAM	Universidad Nacional Autónoma de México
Unesco	Organização das Nações Unidas para a Educação, a Ciência e a Cultura
UNIRIO	Universidade Federal do Estado do Rio de Janeiro
USP	Universidade de São Paulo
XML	Extensible Markup Language

SUMÁRIO

1 INTRODUÇÃO	18
1.1 QUESTÃO DE PESQUISA	21
1.2 HIPÓTESE	21
1.3 OBJETIVOS	21
1.3.1 Objetivo geral.....	21
1.3.2 Objetivos específicos.....	21
1.4 METODOLOGIA.....	22
2 INFORMAÇÃO NO CONTEXTO DA CIÊNCIA DA INFORMAÇÃO	26
2.1 CONCEITUALIZAÇÃO E REPRESENTAÇÃO DA INFORMAÇÃO	27
2.2 ORGANIZAÇÃO DO CONHECIMENTO E ORGANIZAÇÃO DA INFORMAÇÃO ...	30
2.2.1 Sistemas de organização do conhecimento	33
<i>2.2.1.1 Esquemas de classificação</i>	<i>35</i>
<i>2.2.1.2 Vocabulários controlados e tesouros.....</i>	<i>42</i>
<i>2.2.1.3 Códigos de catalogação</i>	<i>46</i>
<i>2.2.1.4 Ontologias</i>	<i>49</i>
2.3 ORGANIZAÇÃO DA INFORMAÇÃO	51
2.3.1 Catalogação	52
2.3.2 Indexação e catalogação de assunto	53
3 INFORMAÇÃO NO CONTEXTO DO JORNALISMO, A NOTÍCIA.....	58
3.1 AS NOTÍCIAS IMPRESSAS, O JORNAL CONTEMPORÂNEO	61
4 DA NECESSIDADE À RECUPERAÇÃO DA INFORMAÇÃO	67
4.1 CONTEXTO NA CIÊNCIA DA INFORMAÇÃO.....	70
4.2 BUSCA DA INFORMAÇÃO E RECUPERAÇÃO DA INFORMAÇÃO	72
4.3 SISTEMAS DE RECUPERAÇÃO DA INFORMAÇÃO.....	76
4.4 LÓGICA E ESTRATÉGIA DE BUSCAS	81
4.5 COMPETÊNCIA EM INFORMAÇÃO	87
5 BIBLIOTECAS.....	91
5.1 BIBLIOTECAS NACIONAIS	92
5.2 BIBLIOTECAS DIGITAIS	93
5.2.1 Digitalização de acervos e o uso do OCR	95
5.3 HEMEROTECAS.....	100
6. BIBLIOTECAS NACIONAIS COM HEMEROTECAS DIGITAIS	104

6.1 ARGENTINA: TRAPALANDA	105
6.2 ÁUSTRIA: ANNO	109
6.3 AUSTRÁLIA: TROVE	116
6.4 BÉLGICA: BELGICAPRESS.....	122
6.5 COSTA RICA	126
6.6 ESPANHA: HEMEROTECA DIGITAL	129
6.7 ESTADOS UNIDOS: <i>CHRONICLING AMERICA</i>	133
6.8 FRANÇA: GALLICA	137
6.9 LIECHTENSTEIN: ELIECHTENSTEINENSIA	141
6.10 MÉXICO: HEMEROTECA NACIONAL DIGITAL DE MÉXICO.....	143
6.11 NOVA ZELÂNDIA: PAPERS PAST	147
6.12 REINO UNIDO: BRITISH NEWSPAPER ARCHIVE	151
6.13 RESUMO DAS CARACTERÍSTICAS DAS HEMEROTECAS ESTRANGEIRAS ..	154
7 BRASIL: DO PRIMEIRO JORNAL À HEMEROTECA DIGITAL.....	169
7.1 BIBLIOTECA NACIONAL DO BRASIL	184
7.2 BIBLIOTECA DIGITAL DA BIBLIOTECA NACIONAL DO BRASIL	188
7.3 HEMEROTECA DIGITAL BRASILEIRA.....	190
8 BUSCA NA HEMEROTECA DIGITAL BRASILEIRA	194
8.1 NECESSIDADES DE INFORMAÇÃO DOS PESQUISADORES	194
8.2 PROCESSOS DE BUSCA E RESULTADOS.....	198
9 INTERPRETAÇÃO DOS RESULTADOS.....	213
10 CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES	226
REFERÊNCIAS	229

1 INTRODUÇÃO

Bibliotecas Nacionais são espaços de preservação da memória e responsáveis pela divulgação e propagação do patrimônio bibliográfico de seus respectivos países. Desde a década de 1990, uma das estratégias de desenvolvimento adotadas por essas instituições tem sido a digitalização de suas coleções, tanto para o próprio processo de preservação quanto para facilitar o acesso e a disseminação dos conteúdos guardados em seus acervos.

A iniciativa resultou na criação de bibliotecas digitais, espaços dinâmicos, constituídos de informações eletrônicas e serviços informatizados que possibilitam inúmeras formas de disseminação e uma gama extraordinária de usos e reusos variados de seus estoques informacionais, conforme ressalta Sayão (2009, p. 14).

Somente a digitalização de acervos, porém, sem o aperfeiçoamento dos processos de tratamento da informação diretamente relacionados (como organização, indexação, sistemas de busca e recuperação da informação e capacitação do usuário), mostra-se insuficiente para a salvaguarda de um acervo, assim como não garante o acesso à informação contida nos documentos armazenados.

É necessário ir além, estabelecer padrões, procedimentos e processos que abranjam a transferência da informação de um suporte para outro, do papel para o digital, e aprimorem o acesso ao conteúdo. Como afirmava Souza já em 2007 (p. 120), os novos ambientes de compartilhamento de conhecimento exigem um repensar de instrumentos e métodos de intermediação necessários para produção, organização e disseminação da informação.

A documentação tradicional, guardada em instituições de preservação da memória, cedeu lugar a uma nova realidade virtual em permanente construção, capaz de se transferir de lugar e mídia em fração de segundos, reproduzindo-se ilimitada e simultaneamente no tempo e no espaço (BETTENCOURT, 2014, p. 17).

A forma de acesso à informação sofreu alterações significativas nas últimas décadas: a maneira indireta de se chegar ao documento e ao seu conteúdo intelectual, por meio da ajuda de um funcionário da unidade de informação, passou a ser direta, feita pelo próprio usuário, em tempo real e por meio de uma rede mundial de comunicação, sendo a “Web” o melhor exemplo dessas novas práticas.

Para atender às necessidades surgidas nesse novo momento, estudos sobre interoperabilidade de sistemas, organização e representação da informação, recursos e ferramentas de busca e recuperação da informação ganharam importância e espaço nos institutos de pesquisa, especialmente apoiados por investigações da Ciência da Informação. Afinal, uma das aspirações dos investigadores dessa área é diminuir a distância entre o usuário e o sistema de informação.

Inserido nesse cenário, este trabalho tem como objeto de pesquisa as hemerotecas, ou seja, coleções de jornais, revistas e/ou publicações periódicas, mais especificamente as hemerotecas digitais de Bibliotecas Nacionais criadas a partir da digitalização de jornais e revistas antigos. Busca-se, por meio da análise de diferentes hemerotecas digitais, identificar os alcances e as limitações dessas unidades de informação.

Para isso, estuda-se especialmente a Hemeroteca Digital Brasileira (HDB) da Biblioteca Nacional (BN), mas também outros 12 projetos, os quais serão apresentados no decorrer das próximas seções, procurando-se, ao final, traçar um retrato do atual estágio de desenvolvimento das hemerotecas digitais de Bibliotecas Nacionais.

Parte significativa das coleções de jornais e revistas em Bibliotecas Nacionais ganharam versões digitais nos últimos 15 anos. O trabalho de digitalização está longe do fim, mas acervos robustos já podem ser acessados nas bibliotecas – e hemerotecas – digitais de diversas instituições.

Dificuldades técnicas de armazenamento parecem ter sido superadas ou pelo menos amenizadas com o avanço de tecnologias, equipamentos e padrões de digitalização e com o barateamento dos custos de estocagem dessa informação em formato digital – apesar do enorme desafio de manter *softwares* e *hardwares* sempre atualizados.

O perfil dos usuários de hemerotecas tende a ser diversificado uma vez que jornais abordam os mais variados assuntos. Assim, coleções armazenadas nesses repositórios são utilizadas para sanar a curiosidade sobre determinado episódio ocorrido em uma cidade ou em um país, ser uma opção de passatempo ou ainda servir de base para pesquisas sobre assuntos específicos em um determinado período, como se pode perceber nos comentários enviados à Biblioteca Nacional do Brasil por usuários da HDB.

Desde o seu lançamento, esse instrumento tem sido a minha maior alegria na internet e tem me ajudado de forma inimaginável, como – tenho certeza – deve estar ajudando também a muitos outros pesquisadores. Graças a vocês pude reconstruir documentalmente a trajetória de uma artista cuja memória estaria perdida e espremida entre jornais velhos não fossem as facilidades que essa página trouxe à pesquisa. Assim, pude fazer justiça à memória de um nome outrora popular e hoje apagado pelo tempo, Georgina Pires Sampaio, a Suzi King.

Pesquisando sobre meu avô materno, Sr. Pedro Argemiro Dias, apareceu uma publicação com o nome dos meus avós maternos (Pedro Argemiro Dias e Isaltina Vieira de Moraes). Gostaria de saber mais, pois meu avô, Pedro Argemiro Dias assassinou uma filha e foi condenado a ficar no Juquery. Fugiu de lá e foi encontrado morto. Como faço pra conhecer a história do meu avô? Minha mãe ainda é viva e tem 90 anos, está lúcida.

Boa disponibilização de material jornalístico. Pra quem gosta de pesquisar o Rio Antigo, como eu, é uma grata surpresa.

Adorei o acesso aos documentos valiosíssimos para nós historiadores.

Fico feliz em saber [da HDB]. Fui um frequente pesquisador na Biblioteca Nacional nos anos 1950. Melhor se puder, agora, fazer isso de casa.¹

Mais do que atender à curiosidade de um público que se interessa por temas gerais, pesquisas familiares e assuntos aleatórios presentes nas páginas de jornais e revistas, os registros passíveis de serem recuperados nesses materiais apresentam-se como preciosa fonte de informação no cenário acadêmico. Afinal, a informação jornalística é extraída do cotidiano, sendo um recorte espaço-temporal da realidade, mesmo que uma realidade parcial, figurando muitas vezes como fonte informacional única.

Há cerca de quatro décadas, cientistas passaram a lançar novo olhar às publicações periódicas, a considerá-las como insumo para suas investigações. Por isso, também com o intuito de chamar a atenção para esta nova possibilidade de acesso e uso desses acervos, este trabalho toma emprestado assuntos e inquietações de um público especializado. Procura-se mostrar o potencial do jornal como opção para buscar um ponto de vista a mais, sendo este, o jornal, um produto de diversos acontecimentos, signos e significados, ou seja, um retrato do dia a dia capaz de preservar a memória nacional, local e institucional.

Neste estudo, a fim de conectar a ciência aos jornais, são apresentadas buscas realizadas na HDB a partir de necessidades reais de pesquisadores envolvidos em investigações científicas. Observa-se a partir desses exemplos que a recente oferta *on-line* de coleções de jornais e revistas pelas Bibliotecas Nacionais facilitou e solidificou o uso de notícias como matéria-prima em pesquisas científicas, seja como elemento principal, complementar ou produzindo *insights*. Por meio deste trabalho, espera-se despertar o interesse e a curiosidade de outros cientistas para estas preciosas fontes de informação que são as hemerotecas digitais.

É inegável a importância da digitalização de acervos antigos para facilitar o acesso a informações que antes só poderiam ser consultadas em endereços físicos únicos e nem sempre de fácil ingresso a todos os usuários interessados. Além disso, a digitalização colabora para a preservação dos originais à medida que os usuários podem recorrer à cópia em versão digital e não mais àquela em papel. Em relação às coleções de jornais e revistas, em muitos casos, os exemplares guardados nas Bibliotecas Nacionais são cópias únicas, representando um verdadeiro tesouro da memória de uma determinada época.

Em resumo, oferecer coleções de jornais em formato digital e disponível na “Web” reduz a necessidade de deslocamento dos usuários, aumenta o interesse por edições antes de

¹ Depoimentos cedidos pela Hemeroteca Digital Brasileira (2012).

difícil acesso e permite a pesquisa no momento mais adequado ao indivíduo, podendo este escolher quando e de onde acessar o acervo digitalizado.

Todavia, os problemas relacionados a organização, representação e recuperação da informação persistem. São temas instigantes de investigação que se apresentam aos cientistas da informação dispostos a aproximar o usuário das respostas que melhor atendam as suas necessidades informacionais.

1.1 QUESTÃO DE PESQUISA

Quais são os alcances e os limites das hemerotecas digitais criadas por Bibliotecas Nacionais na oferta de informações que possam servir de apoio à pesquisa científica?

1.2 HIPÓTESE

Jornais e revistas guardam, em notícias, reportagens, crônicas e comentários de leitores e propagandas, informações valiosas àqueles pesquisadores interessados em recontar a história da sociedade a partir da realidade esboçada nos registros não oficiais escritos por jornalistas nas páginas das antigas publicações agora acessíveis nas hemerotecas digitais de Bibliotecas Nacionais, e que podem contribuir para o seu trabalho de investigação.

O investimento realizado pelas BN na criação de hemerotecas digitais tem permitido o acesso facilitado ao acervo de jornais e revistas. A consulta *on-line* evita deslocamentos geográficos, preserva a coleção e representa economia de tempo e dinheiro. Além disso, não é mais necessário procurar manualmente pela informação desejada, ela pode ser recuperada com certa precisão por meio de novos recursos.

1.3 OBJETIVOS

1.3.1 Objetivo geral

Estudar características de hemerotecas digitais criadas por Bibliotecas Nacionais, compondo um retrato do *status* de desenvolvimento dessas coleções, a fim de buscar subsídios para esboçar um modelo de conhecimento que possibilite a pesquisadores científicos obter melhores resultados ao realizar buscas em jornais e revistas antigos.

1.3.2 Objetivos específicos

1. Identificar procedimentos, padrões, melhores práticas e recursos de hemerotecas digitais de Bibliotecas Nacionais ligados ao uso de acervos de jornais e revistas antigos.

2. Examinar a potencial utilidade da HDB a pesquisadores após relatar a trajetória da Hemeroteca Digital Brasileira e de realizar buscas em seu acervo com base nas necessidades de informação de pós-graduandos em História das Ciências e da Saúde.
3. Por meio dos resultados de uso da HDB, embasar aprimoramentos de recursos de informação e estimular pesquisadores científicos à reflexão sobre o planejamento da definição da necessidade de informação e do processo de busca em coleções de jornais e revistas.

1.4 METODOLOGIA

A metodologia de uma pesquisa científica determina o roteiro de ações programadas e coordenadas que conduz ao desenvolvimento de algo novo. González de Gómez (2000, p. 1) afirma que “a metodologia de pesquisa designa o início e a orientação de um movimento de pensamento cujos esforço e intenção direcionam-se à produção de um novo conhecimento, em um horizonte de possibilidades sociais e historicamente definidas”.

Esta pesquisa exploratória iniciou-se com a revisão do conceito de informação nos âmbitos da Ciência da Informação e do Jornalismo. Não apenas por ser a informação o elemento base das duas áreas, mas por estarem a Ciência da Informação e o Jornalismo intimamente ligados ao objeto de estudo desta tese: as hemerotecas digitais criadas por Bibliotecas Digitais.

No domínio da Ciência da Informação, por meio da revisão bibliográfica buscou-se as definições de organização do conhecimento e de organização da informação, embasamento que se mostrou necessário para compreender seus desdobramentos, ou seja, a representação do conhecimento e a representação da informação. A primeira mostrando-se como suporte para a concretização da segunda.

Nesta etapa da revisão, relatar a trajetória e a evolução dos sistemas de organização do conhecimento mostrou-se relevante para a compreensão do trabalho desenvolvido desde os primórdios da Biblioteconomia e da Ciência da Informação por ilustrar o grande esforço empreendido ao longo de décadas para tornar mais organizada e acessível a informação armazenada em unidades de informação.

Apresentados os mecanismos desenvolvidos para raciocinar sobre o mundo, o foco recaiu na aplicação prática desses esquemas de representação do conhecimento, sobre como passaram a ser utilizados no dia a dia de uma unidade de informação. O endereçamento das publicações nas estantes de uma biblioteca a partir da Classificação Decimal Universal, a indexação de projetos científicos baseada em um vocabulário controlado e o desenvolvimento

do catálogo de uma hemeroteca por meio do uso da AACR2.

Ainda no âmbito da Ciência da Informação, buscou-se embasamento teórico para compreender os processos de busca da informação e de recuperação da informação, assim como a formulação da necessidade de informação, os mecanismos e as estratégias que permitam ao usuário chegar mais próximo das respostas que procura nos acervos de unidades de informação. Para entender como este usuário avalia – ou deveria avaliar – os resultados, reuniram-se noções sobre competência em informação.

No âmbito do Jornalismo foram abordados o jornal, sua trajetória e características, a informação jornalística, a notícia e o jornal como fonte histórica. Antes de iniciar a apresentação das coleções armazenadas, digitalizadas e disponibilizadas pelas Bibliotecas Nacionais, foi revista a trajetória e o significado de biblioteca, biblioteca digital, hemerotecas, assim como repassado o processo de digitalização de acervos e o uso do recurso de reconhecimento óptico de caracteres (OCR).

Em seguida, iniciou-se a etapa empírica da pesquisa, com a seleção das hemerotecas digitais. Tendo como referência a lista de países integrantes da Organização das Nações Unidas (ONU), foram realizadas buscas no *Google* a fim de localizar as páginas das Bibliotecas Nacionais e suas respectivas hemerotecas. Esta etapa ocorreu em março de 2015.

Ao iniciar a seleção das hemerotecas digitais de BN, estabeleceram-se dois critérios principais:

- a) Existência de uma hemeroteca digital, ou seja, um acervo digitalizado de jornais e revistas que pudesse ser consultado via “Web”;
- b) A hemeroteca estar disponível em algum dos seguintes idiomas: português, inglês, espanhol, alemão, francês ou italiano. Todos os sites com página inicial, mecanismo de busca ou acervo em idiomas diferentes desses foram descartados devido a limitações linguísticas.

Assim, dos 193 países pesquisados, observou-se que 108 mantinham sites com informações sobre as respectivas Bibliotecas Nacionais. Porém, por não atenderem aos dois critérios mencionados, 95 foram descartados. Em dez casos, como o da Alemanha, a Biblioteca Nacional mantinha uma biblioteca digital, mas não uma hemeroteca digital.

As treze hemerotecas digitais selecionadas foram:

1. Trapalanda, da Biblioteca Nacional da Argentina;
2. Trove, da Biblioteca Nacional da Austrália;
3. Anno, da Biblioteca Nacional Austríaca;
4. BelgicaPress, da Biblioteca Nacional da Bélgica;

5. Hemeroteca Digital Brasileira, da Biblioteca Nacional do Brasil;
6. Hemeroteca Digital da Biblioteca Nacional da Costa Rica;
7. Hemeroteca Digital da Biblioteca Nacional da Espanha;
8. Chronicling America, da Biblioteca do Congresso dos Estados Unidos;
9. Gallica, da Biblioteca Nacional da França;
10. eLiechtensteinesia, da Biblioteca Nacional de Liechtenstein;
11. Hemeroteca Nacional Digital da Biblioteca Nacional do México;
12. Papers past, da Biblioteca Nacional da Nova Zelândia;
13. British Newspaper Archive, da Biblioteca Nacional Britânica.

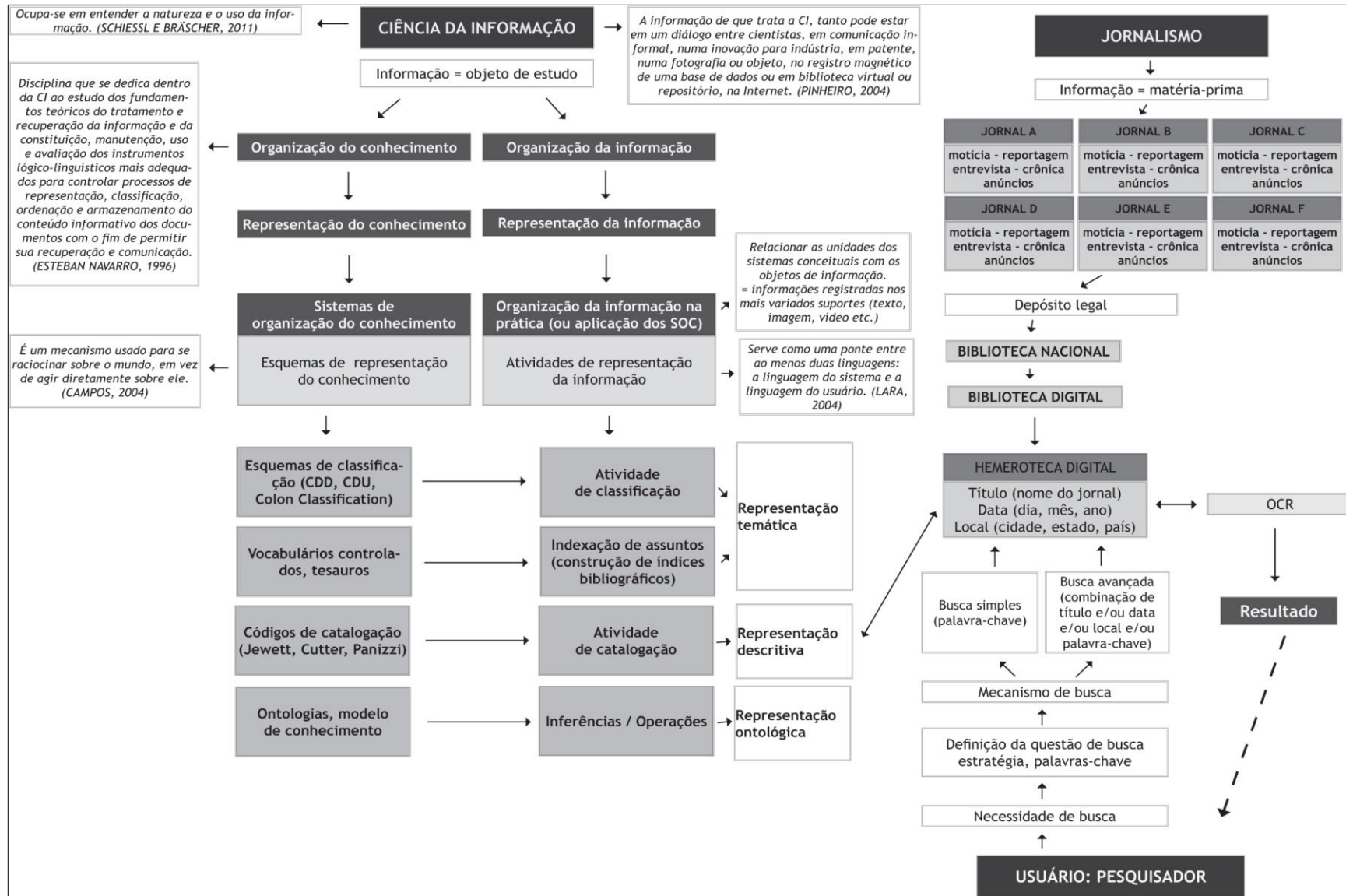
O passo seguinte englobou pesquisa sobre as hemerotecas digitais selecionadas, exceto a Hemeroteca Digital Brasileira, tratada em separado.

As informações sobre as hemerotecas foram obtidas por meio da navegação nos referidos sites e via entrevista semiestruturada realizada por e-mail. Das doze hemerotecas digitais envolvidas, oito responderam com informações sobre diferentes aspectos, como trajetória da formação da hemeroteca digital, critérios de seleção do acervo, recursos e ferramentas de armazenamento, organização e mecanismos de recuperação das informações, tamanho do acervo, planejamento, execução e desdobramentos do projeto e opções de resultados disponíveis. O material coletado foi inserido nos textos das respectivas hemerotecas.

Realizou-se então um aprofundamento sobre a Hemeroteca Digital Brasileira englobando a história da imprensa no Brasil, a trajetória da Biblioteca Nacional do Brasil, o surgimento da BNDigital até chegar à criação da Hemeroteca Digital Brasileira (HDB).

Por fim, as possibilidades de uso da coleção de jornais digitalizados pela Biblioteca Nacional brasileira foram testadas em buscas na HDB, que tiveram como base as informações coletadas por meio de entrevistas semiestruturadas, no período de 15 a 30 de outubro de 2015, aplicadas a pesquisadores da área de História das Ciências e da Saúde, pós-graduandos ou egressos do Programa em História das Ciências e da Saúde (PPGHCS) da Casa de Oswaldo Cruz (COC), pertencente à Fundação Oswaldo Cruz (Fiocruz), no Rio de Janeiro, e a um mestrando em História da Universidade Federal de Santa Catarina (UFSC), em Florianópolis. A área dos pesquisadores foi escolhida por ser bastante abrangente, contemplando variados temas, o que tornaria a busca de informação mais diversificada. A partir das buscas de informação empreendidas, obteve-se subsídios para se esboçar o pretendido modelo de conhecimento. A trajetória percorrida na revisão bibliográfica é apresentada na Figura 1.

Figura 1 – Percurso da revisão bibliográfica



2 INFORMAÇÃO NO CONTEXTO DA CIÊNCIA DA INFORMAÇÃO

“Informação é aquilo que alimenta o funcionamento do nosso mundo: o sangue e o combustível, o princípio vital. A informação permeia a ciência de cima abaixo, transformando todos os ramos do conhecimento.” James Gleick (2013) utiliza mais de quatro centenas de páginas para descrever esta palavra de difícil definição, presente nos mais diversos cenários, sendo elemento base tanto da Ciência da Informação quanto do Jornalismo.

Polissêmica e considerada como recurso básico da atividade humana e, por consequência, de diferentes áreas do conhecimento, informação só pode ser explicada e esmiuçada dentro de contextos bem determinados – fora disso, corre o risco de assemelhar-se a uma obra de M.C. Escher, como observa Morville (2005, p. 45). Nesta revisão de literatura serão considerados os significados do termo informação nos domínios da Ciência da Informação e do Jornalismo.

Morville (2005, p. 46) afirma que informação está intrinsecamente ligada à comunicação, envolvendo a troca de símbolos, ideias, mensagens e significados entre as pessoas. Sendo assim, é caracterizada pela ambiguidade, redundância, ineficiência, erro e indescritível beleza. A comunicação está entre as maiores capacidades da humanidade, sendo a sua espinha dorsal. Além disso, a informação permite às pessoas se comunicarem através do tempo e do espaço. Exemplo disso é o acesso que se tem hoje a notícias publicadas em jornais décadas atrás nos mais diferentes lugares do mundo.

A informação comporta um elemento de sentido. É um significado transmitido a um ser consciente por meio de uma mensagem inscrita em um suporte espacial-temporal: impresso, sinal elétrico, onda sonora etc. Essa inscrição é feita graças a um sistema de signos (linguagem), signo este que é um elemento da linguagem que associa um significante a um significado: signo alfabético, palavra, sinal de pontuação (LE COADIC, 2004, p. 5).

Seja na Ciência da Informação ou no Jornalismo, quando um indivíduo busca uma informação, ele está de fato à procura de um conhecimento que ainda não detém. Conhecimento, segundo Le Coadic (2004, p. 5), é o resultado do ato de conhecer, ou seja, de ser capaz de formar a ideia de alguma coisa. Em seus escritos sobre a história social do conhecimento, Burke (2003, p. 19) adota uma maneira prática de definir e diferenciar informação e conhecimento. O autor usa o termo informação para referir-se ao que é relativamente “cru”, específico e prático, e conhecimento para denotar o que foi “cozido”, processado ou sistematizado pelo pensamento.

Lima e Alvares (2012, p. 24) escrevem que o conhecimento está relacionado com os aspectos cognitivos que ocorrem na mente humana e envolvem os processos mentais de captação, assimilação, associação e de construção, desconstrução e reconstrução de conceitos.

A comunicação efetiva do conhecimento e de seus registros entre os seres humanos, no contexto social, institucional ou individual do uso e das necessidades de informação, figura entre os assuntos tratados pelos cientistas da informação, como destaca Saracevic (1996, p. 47). Schiessl e Bräscher (2011, p. 301) afirmam que “a Ciência da Informação se ocupa em entender a natureza e o uso da informação”.

Acerca da informação na Ciência da Informação, Pinheiro destaca que:

Informação é tradicionalmente relacionada a documentos impressos e a bibliotecas, quando de fato a informação de que trata a Ciência da Informação, tanto pode estar em um diálogo entre cientistas, em uma comunicação informal, numa inovação para indústria, em uma patente, numa fotografia ou em um objeto, no registro magnético de uma base de dados ou em uma biblioteca virtual ou repositório, na Internet (PINHEIRO, 2004, p. 1).

Capurro e Hjørland (2007, p. 187), após debruçarem-se de forma exaustiva sobre o termo informação, coletando definições e pontos de vista, afirmam que esta informação de que trata a Ciência da Informação “é qualquer coisa de importância na resposta a uma questão”, ou seja, “informação é o que pode responder questões importantes relacionadas às atividades do grupo-alvo”.

Um grupo-alvo, na Ciência da Informação, pode ser formado por usuários de diversas áreas, com necessidades específicas de informação e noções variadas do que é relevante. O atendimento ao que necessitam esses indivíduos, que se inicia na representação da informação, apresenta-se como o *Leitmotiv* dos profissionais da informação – e seu maior desafio.

2.1 CONCEITUALIZAÇÃO E REPRESENTAÇÃO DA INFORMAÇÃO

Na Ciência da Informação, para que uma informação torne-se recuperável é necessário representá-la, ou seja, como definem Lima e Alvares (2012, p. 21), “utilizar elementos simbólicos – palavras, figuras, imagens, desenhos, mímicas, esquemas, entre outros – para substituir um objeto, uma ideia ou um fato”. Representar ou descrever é algo realizado pelo homem desde tempos remotos, quando registrava itens pictográficos nas paredes de cavernas ou controlava o tamanho de suas posses por meio de rabiscos em tabletes de argila.

Ao gravar informações em barro, madeira ou pedra, o homem deu importante passo no processo civilizatório, contribuindo para a preservação do conhecimento humano. Das inscrições rupestres, o indivíduo evoluiu para a escrita, “organização de sinais gráficos em linhas”, como a define Flusser (2010, p. 28). O surgimento da escrita ocorreu de forma diacrônica, mas os registros tidos como os mais antigos foram encontrados no chamado Crescente Fértil, área das primeiras grandes civilizações.

Na Mesopotâmia, entre os rios Tigre e Eufrates, os sumérios desenvolveram a escrita cuneiforme para representar, especialmente, informações contábeis e administrativas. “As tabuletas não apenas registravam o comércio e a burocracia, mas, antes de mais nada, os possibilitaram”, afirma Gleick (2013, p. 51).

A escrita dos sumérios foi adotada, posteriormente, por outras civilizações na região, como as dos babilônios e assírios. Às margens do rio Nilo, egípcios desenvolveram hieróglifos, hierática e demótico. “De repente, algo intangível – um número, uma notícia, um pensamento, uma ordem – podia ser obtido sem a presença física do mensageiro; magicamente, podia ser imaginado, anotado e passado adiante no espaço e no tempo” (MANGUEL, 2004, p. 100).

A escrita surge em consequência da necessidade de se perenizar o conhecimento e de organizá-lo fora da memória, então insuficiente, face à evolução do próprio homem. A escrita passa a ter uma função não só de armazenar informações e possibilitar a comunicação através do tempo, como também de fornecer ao homem a possibilidade de memorizar, reexaminar, registrar e retificar conteúdos e observar significados (BETTENCOURT, 2014, p. 23).

O alfabeto foi uma tecnologia fundadora da informação. Com o passar dos anos surgiram outras, como afirma Gleick (2013, p. 20): “o telefone, o aparelho de fax, a calculadora e, por fim, o computador são apenas as mais recentes inovações projetadas para armazenar, manipular e transmitir conhecimento”. Para o autor, as tecnologias de informação dificilmente se tornam obsoletas, cada nova tecnologia traz um alento as suas antecessoras. Cada uma serve aos indivíduos pelo tempo que se faz necessário, até que surjam outras ferramentas que atendam de forma mais adequada às necessidades cotidianas.

Não deveríamos desmerecer tecnologias que tomamos hoje como ultrapassadas, pois sem estas talvez não tivéssemos chegado ao estágio atual. Como escreveu Thomas Hobbes, no século XVII, citado por Gleick (2013, p. 20), “a invenção da imprensa é insignificante se comparada à invenção das letras”.

Darnton (2010, p. 41) observa que cada mudança na tecnologia transformou o panorama da informação. O autor argumenta que todas as eras foram eras de informação, cada qual a seu modo. A velocidade das mudanças foi aumentando com o passar do tempo: “da escrita ao códice foram 4300 anos; do códice aos tipos móveis, 1150 anos; dos tipos móveis à internet, 524 anos; da internet aos buscadores, 17 anos; dos buscadores ao algoritmo de relevância do Google, sete anos.” O que mais estará por vir?

Nas primeiras eras da informação, o que o homem representava em seus desenhos, símbolos alfabéticos ou sinais gráficos eram conceitos. Dahlberg (1978b, p. 12) define conceito como unidade do conhecimento, por ser algo externalizado, expresso e comunicado pela linguagem, “não restrito à mente daquele que pensa”, como apontam Francelin e Kobashi (2011, p. 207). Afinal, como a própria Dahlberg (1978a, p. 101) explica, o conhecimento fixa-se por meio de elementos da linguagem, seja de forma oral, seja pela escolha de um termo ou um nome, que será depois registrado.

“A linguagem constitui a capacidade do homem de designar os objetos que o circundam assim como de se comunicar com os semelhantes” (DAHLBERG, 1978a, p. 101). Em necessidades cotidianas, o indivíduo utiliza a chamada linguagem natural, “discurso comum com o qual se comunica no dia a dia, que permite expressar valores, emoções, associação de ideias, enfim, comunicar efetivamente” (SCHIESSL; SHINTAKU, 2012, p. 53).

Como afirmam Guinchat e Menou:

As linguagens naturais, isto é, as linguagens faladas, apresentam características que dificultam sua utilização para o tratamento da informação. Elas são numerosas e têm, cada uma, grande quantidade de termos, o que dificultaria o manuseio dos arquivos. As linguagens naturais são adaptadas a formas de comunicação oral ou escrita, nas quais se estabelece entre os interlocutores uma forma de diálogo. O tempo e o espaço têm papel importante neste diálogo. A linguagem natural pressupõe nuances, associações de ideias, expressão de emoções e de valores (GUINCHAT; MENO, 1994, p. 133).

Nas áreas científicas, como a Biblioteconomia e a Ciência da Informação ou mesmo a Ciência da Computação, o homem usa linguagens especiais, artificiais ou formalizadas para traduzir a linguagem natural, transformando-a em uma linguagem padronizada e estável. Conceitos e termos formam a base da chamada linguagem documentária, cujo “objetivo principal é o controle do vocabulário, ou seja, controlar os termos da área em foco por meio de um conceito/interpretação” (VIEIRA, 2014, p. 65).

É importante distinguir entre conceitos e palavras. Os conceitos são expressos em palavras, mas não são idênticos a palavras. Por exemplo, um inglês usará a palavra *horse* enquanto um francês usa a palavra *cheval* para exatamente o mesmo conceito. É possível também haver um conceito (de alguma coisa) para o qual não haja palavra para expressá-lo ou para o qual não conheçamos a palavra (LANGRIDGE, 2006, p. 21).

Assim como conceito distingue-se de palavra, termo não pode ser usado como sinônimo de palavra. Como afirmam Cesarino e Pinto (1978, p. 268), as palavras não têm significado preciso e apesar disso, na recuperação da informação, é necessário agir como se tivessem. Por essa razão precisa-se reduzir ao máximo o elemento “interpretação pessoal”. Para que possa ser recuperada, esta palavra deverá ser adequadamente identificada ao ser inserida em um sistema. Ela terá de ser transformada em um termo com significado bem estabelecido.

”Termo é uma palavra especializada, ou um conjunto delas, isto é, ele está inserido no contexto das línguas especializadas. As palavras, na língua natural”, definem Schiessl e Shintaku (2012, p. 97), para em seguida completar afirmando que uma palavra pode ter vários significados, mas um termo sempre terá apenas um dentro de um domínio específico.

Com a linguagem, como descreve Dahlberg (1978a, p. 101), o homem tornou-se capaz de se relacionar com os vários itens que o circundavam e de elaborar enunciados sobre tais objetos. Há objetos individuais, presentes em espaço e tempo determinados, no aqui e agora, e há objetos gerais, que prescindem das formas do tempo e do espaço e por isso mesmo podem ser ditos como idealizados. Os primeiros são expressos por conceitos individuais, bem definidos, pertencendo ao mundo real. Os outros, como o nome já indica, são chamados de conceitos gerais e representam construções mentais.

“É fácil verificar que os elementos contidos nos conceitos gerais encontram-se nos conceitos individuais, sendo, portanto, possível reduzir os conceitos individuais aos gerais e ordená-los de acordo com os conceitos gerais” (DAHLBERG, 1978a, p. 102). A formação dos conceitos nada mais é do que a compilação de enunciados verdadeiros sobre determinado objeto, predicados que fornecem uma característica, um elemento.

Tal processo de definição finda ao se escolher um instrumento, ou seja, uma palavra ou um símbolo linguístico, que traduza e fixe a compilação – como os citados no exemplo acima, IBICT e instituição. “A formação dos conceitos é a síntese dos predicados necessários verdadeiros a respeito de determinado objeto” (DAHLBERG, 1978b, p. 12).

A conceituação tem grande importância no trabalho desenvolvido por profissionais da informação. Conceitos são componentes essenciais do conhecimento a ser representado, como aponta Alvarenga (2003, p. 22). A representação é uma ferramenta de trabalho desses profissionais, especialmente dos bibliotecários, que dela se utilizam para tornar recuperáveis os itens de um acervo.

Na próxima seção, a representação será abordada nos contextos da organização do conhecimento e da organização da informação.

2.2 ORGANIZAÇÃO DO CONHECIMENTO E ORGANIZAÇÃO DA INFORMAÇÃO

Representar é reformular ou indicar um conceito por meio de uma linguagem de indexação, como definido no *Dicionário de Biblioteconomia e Arquivologia* (CUNHA; CAVALCANTI, 2008, p. 322). Tal tarefa está intrinsecamente ligada à atuação dos profissionais da informação, cujo papel mais importante é desenvolver meios e ferramentas capazes de facilitar o acesso dos usuários à informação que atenderá as suas necessidades informacionais.

Como afirma Biolchini:

A eficácia dos processos de representação da informação reside na medida em que os elementos contidos no estoque informacional encontram-se "perceptíveis" pelo usuário através do sistema, de forma que o modo de organização da informação contida no estoque apresente condições de acessibilidade, otimizada em função dos meios utilizados para sua representação (BIOLCHINI, 2001).

No âmbito da Ciência da Informação, a representação está relacionada tanto à informação quanto ao conhecimento. A representação da informação está inserida nos estudos de organização da informação. A representação do conhecimento faz parte dos estudos de organização do conhecimento.

Nesta revisão bibliográfica obteve-se acesso a um emaranhado de definições e crenças. As duas abordagens são consideradas como sinônimos por determinados autores, mas neste estudo será considerada a versão adotada pelo grupo de estudiosos que defende se tratem de dois enfoques de pesquisa distintos.

Desta forma, os conceitos estão ligados à organização do conhecimento (OC), processo que “visa à construção de modelos de mundo que se constituem em abstrações da realidade”, como afirmam Bräscher e Café (2010, p. 6).

Para Esteban Navarro,

OC é a disciplina específica que se dedica dentro da Ciência da Informação ao estudo dos fundamentos teóricos do tratamento e recuperação da informação e da constituição, manutenção, uso e avaliação dos instrumentos lógico-linguísticos mais adequados para controlar processos de representação, classificação, ordenação e armazenamento do conteúdo informativo dos documentos com o fim de permitir sua recuperação e comunicação. Ocupa-se, portanto, dos princípios e ferramentas postos em ação para a gestão do conhecimento humano a partir de uma perspectiva tripla: sua representação, sua organização e sua comunicação documental (ESTEBAN NAVARRO, 1996, p. 97-98, tradução nossa).

Assim, a OC, no escopo da Ciência da Informação, é a área de estudos com foco nas atividades de organização, representação e recuperação da informação. “Representar o conhecimento é uma tentativa de se apropriar dos elementos informacionais existentes nas estruturas e processos mentais que compõem o conhecimento individual, para que o saber possa ser socializado” (LIMA; ALVARES, 2012, p. 33). Para Campos (2004, p. 24), “uma representação de conhecimento é um mecanismo usado para se raciocinar sobre o mundo, em vez de agir diretamente sobre ele”.

O ponto de partida para o conhecimento é a informação. Somente após obter uma informação, o indivíduo terá matéria-prima para transformá-la em conhecimento. O processo, porém, não é automático. Nesse processo de assimilação e transformação da informação em conhecimento também está inserida a questão da competência em informação. Para que a informação obtida por um indivíduo se transforme em conhecimento, ela precisa dialogar com

a cultura desse sujeito, seus valores e princípios, com os saberes prévios armazenados em sua mente, assim como seu modo de ser e a maneira como se relaciona e compreende o mundo.

Borko (1968, p. 4) destaca que a pesquisa em Ciência da Informação investiga as propriedades e o comportamento da informação, a utilização e a transmissão da informação, bem como o processamento da informação para ótima armazenagem e recuperação. Assim, elaborar modelos e investigar quais as melhores formas de representar determinada informação para que seja recuperada de maneira mais precisa confirma-se como uma das atividades centrais da Ciência da Informação.

Se na OC, a questão é desenvolver metodologias e construir sistemas conceituais – Sistemas de Organização do Conhecimento (SOC) –, na organização da informação (OI), a preocupação está em relacionar as unidades dos sistemas conceituais com os itens que compõem uma unidade da informação. Isso se dá por meio dos recursos desenvolvidos na OC. A partir dessas ferramentas é possível criar representações descritivas e temáticas dos objetos. Para Lima e Alvares (2012, p. 34), no domínio da OI estão os meios para recuperar a informação com vistas ao melhor desempenho e fidedignidade.

Em um sistema de informação, a qualidade obtida na recuperação da informação depende substancialmente dos procedimentos e instrumentos utilizados na OI. [...] O principal objetivo da OI é recuperar objetos informacionais, que são as informações registradas nos mais variados suportes – textos, imagens, registros sonoros, representações cartográficas, páginas “Web”, entre outros (LIMA; ALVARES, 2012, p. 35).

Como referido, a OI é uma questão central para que o processo de recuperação da informação atenda à necessidade do usuário. Quando a informação não está organizada (isto é, representada de acordo com modelos pré-estabelecidos), é difícil, se não impossível, encontrá-la e, conseqüentemente, recuperá-la. “O processo de recuperação da informação depende muito do tratamento e do armazenamento da informação, que, em grande medida, determinam as melhores estratégias para as buscas em um Sistema de Recuperação da Informação”, afirma Miranda (2005, p. 62).

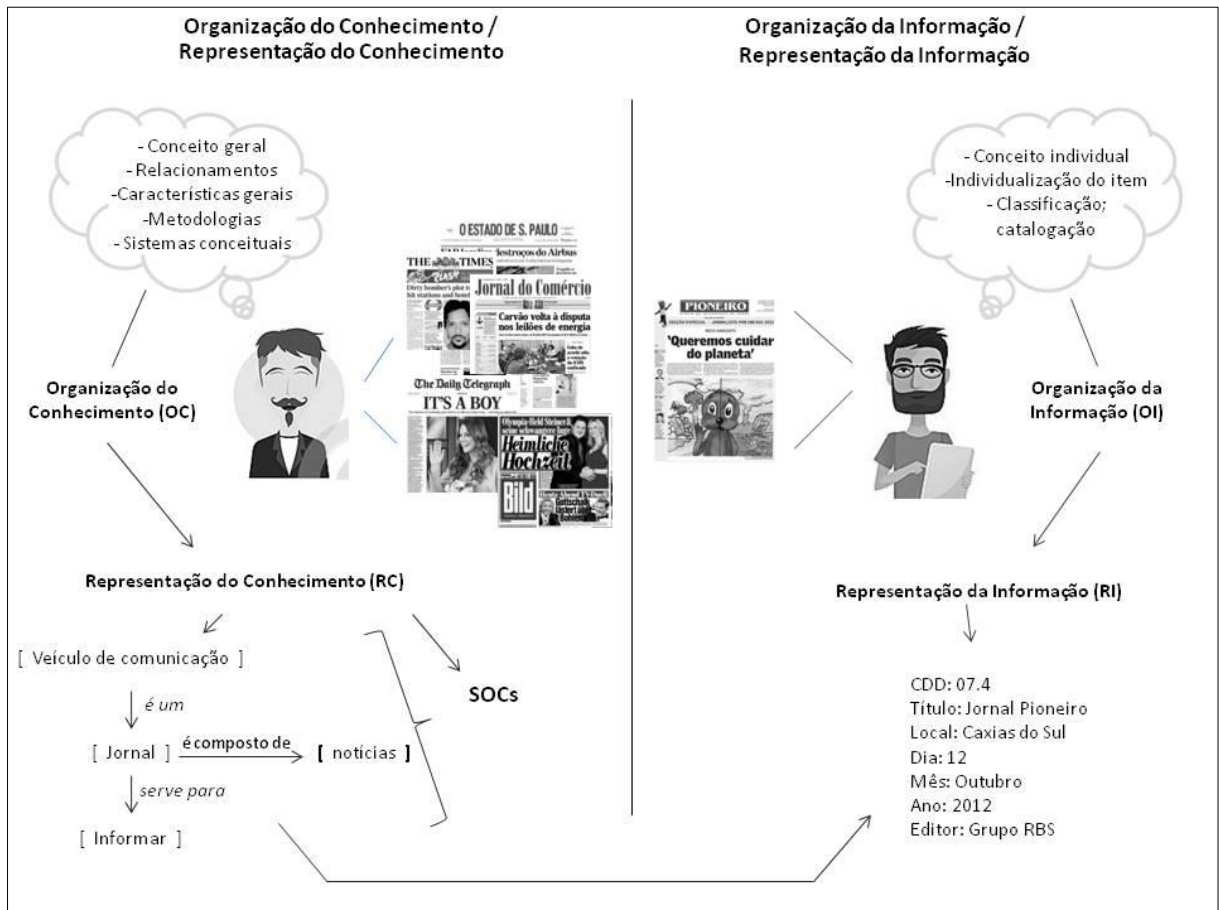
“Ao fazer uma busca em uma base de dados, o que se almeja é encontrar documentos que sejam úteis para satisfazer uma necessidade de informação”, avalia Lancaster (2004, p. 4).

Para Bräscher e Café,

OI é, portanto, um processo que envolve a descrição física (do suporte da informação) e de conteúdo (conhecimento) dos objetos informacionais. O produto desse processo descritivo é a representação da informação (RI), entendida como um conjunto de elementos descritivos que representam os atributos de um objeto informacional específico (BRÄSCHER; CAFÉ, 2010, p. 5).

A partir do modelo desenvolvido por Bräscher e Café (2010, p. 5), esquematizou-se a OC/RC e a OI/RI (Figura 2).

Figura 2 – OC/RC, OI/RI



Fonte: A autora (2016) a partir de modelo de Bräscher e Café (2010, p. 5).

Bräscher e Café reforçam ainda que, no contexto da OI e da RI, objeto de estudo são os registros de informação. Trata-se então do mundo físico, diferente do mundo da cognição e das ideias, cuja unidade elementar é o conceito – objeto, como dito anteriormente, da OC e da RC, apresentados por meio de sistemas de organização do conhecimento (SOC).

Na seção seguinte, serão abordados os SOC.

2.2.1 Sistemas de organização do conhecimento

A organização do conhecimento compreende sistemas, métodos e técnicas, isto é, ferramentas baseadas em fundamentos teóricos capazes de estruturar e organizar sistematicamente unidades de conhecimento (conceitos). Na prática, a OC materializa-se em sistemas de organização do conhecimento. “SOC são sistemas conceituais semanticamente

estruturados que contemplam termos, definições, relacionamentos e propriedades dos conceitos” (CARLAN; MEDEIROS, 2011, p. 54).

Os SOC incluem a variedade de esquemas que organizam, gerenciam e recuperam a informação. Existem desde os tempos remotos e estão presentes em todas as áreas do conhecimento humano, de modos simples aos mais complexos. Esses sistemas abrangem classificação, tesouro, ontologia, assim como os conhecidos glossários e dicionários específicos a cada área e, em sua maioria, ligados a bibliotecas e outras organizações de gerenciamento da informação visando a organizar, recuperar e disseminar a informação (TRISTÃO; FACHIN; ALARCON, 2004, p. 161).

“Tais sistemas são utilizados no momento do tratamento intelectual dos documentos, isto é, no momento das operações de entrada no subsistema de armazenamento e de pesquisa da informação e no momento das operações de saída e de difusão da informação” (GUINCHAT; MENO, 1994, p. 133). Esses esquemas, como ressaltam Guinchat e Menou, são instrumentos destinados à realização de operações precisas, em condições precisas, para responder a necessidades precisas.

Como define Carlan (2010, p. 16), as funções dos SOC incluem a eliminação de ambiguidade, controlam sinônimos ou equivalentes e estabelecem relacionamentos semânticos entre conceitos. Esta última característica é citada por Vickery (2008) como imprescindível para uma ferramenta ser considerada um SOC. Além disso, escreve Vickery (2008), os sistemas são desenvolvidos para organizar, gerir e recuperar informação.

Os SOC ou esquemas de representação do conhecimento, como alguns autores preferem denominar, são encontrados na literatura das áreas de Ciência da Informação, Biblioteconomia e Documentação para designar instrumentos que fazem a tradução dos conteúdos dos documentos originais e completos, para um esquema estruturado sistematicamente, que representa esse conteúdo, com a finalidade principal de organizar a informação e o conhecimento e, conseqüentemente, facilitar a recuperação das informações contidas nos documentos. A infraestrutura que dá suporte ao desenvolvimento dos SOC requer, antes de mais nada, uma análise das necessidades dos usuários dos sistemas, a identificação do tipo de SOC apropriado e o desenvolvimento do *hardware* e do *software* adequados à arquitetura de rede, sua integração e manutenção (CARLAN, 2010, p. 28).

Os SOC adotam linguagens documentárias, que simplificam e condensam o contexto da linguagem natural para que seja preservada a riqueza contida nos documentos originais. Lara (2004, p. 232) explica que a denominação linguagem documentária designa, de modo mais amplo e completo, a linguagem especialmente construída para organizar e facilitar o acesso e a transferência da informação. “A linguagem documentária serve como uma ponte entre ao menos duas linguagens: a linguagem do sistema e a linguagem do usuário” (LARA, 2004, p. 233).

Schiessl e Shintaku descrevem que:

A linguagem documentária visa representar, de forma padronizada, o conhecimento de domínio específico e estabelecer a ligação entre o usuário e o conteúdo do documento. Ela é construída artificialmente com o propósito de sintetizar por meio de vocabulário controlado, o conteúdo dos documentos, e ser utilizada em sistemas de indexação, armazenamento e recuperação da informação. Enfim, delimita o vocabulário de determinado domínio de interesse de forma clara e precisa (SCHIESSL; SHINTAKU, 2012, p. 54).

Para que a linguagem documentária consiga dar forma ao conteúdo, propondo-se como um modo de organização e desempenhando o papel de instrumento de comunicação, segundo Lara (2004, p. 233), ela precisará reunir determinadas qualidades: funcionar como um código inteligível e como fonte para a interpretação do sentido, assim como caracterizar-se como metalinguagem e incorporar o usuário como integrante do processo, utilizando referências de linguagem, e seu significado, que sejam razoavelmente compartilhadas.

Nos subcapítulos a seguir serão abordados SOC relevantes no escopo deste trabalho: sistemas de classificação, vocabulários controlados, tesouros, ontologias e recursos para a catalogação.

2.2.1.1 Esquemas de classificação

Logo nas primeiras páginas de seu fundamental *Classificação: abordagem para estudantes de Biblioteconomia*, Langridge destaca a natureza exemplar do processo de classificar. “Sem classificação não poderia haver o pensamento humano, a ação e a organização que conhecemos”, afirma Langridge (2006, p. 11), para em seguida citar John Dewey, autor da frase “conhecimento é classificação”. Sobre a qual Langridge comenta talvez não ser uma definição muito precisa de conhecimento, mas que enfatiza o papel significativo da classificação.

Classificar, como define Piedade (1983, p. 16) é dividir em grupos ou classes, levando em conta semelhanças e diferenças em certo número de grupos metodicamente distribuídos. A autora cita Ranganathan, para quem classificar consistia em traduzir o nome dos assuntos dos documentos da linguagem natural para a linguagem artificial dos sistemas de classificação bibliográfica.

As semelhanças e as diferenças são observadas nos atributos do que está sendo classificado, ou seja, nas características, que podem ser naturais ou artificiais. “É natural quando é inerente ou inseparável do objeto a classificar. É artificial quando é ocasional, acidental e variável” (PIEADADE, 1983, p. 18). A partir das características, objetos ou ideias são agrupados em classes que se enquadram em categorias, ou seja, classes mais gerais. Exemplo clássico de categorias são aquelas estabelecidas por Ranganathan: personalidade,

matéria, energia, espaço e tempo (PMEST).

De acordo com Guinchat e Menou (1994, p. 142), as classificações são as linguagens documentais mais antigas, e são adaptadas especialmente ao trabalho das bibliotecas – de certa forma, a classificação é o endereço físico do item. Existem vários tipos, sendo que as classificações universais ou enciclopédicas cobrem todos os ramos do conhecimento.

Os sistemas de classificação podem ser dispostos de diferentes formas, segundo suas qualidades, características, compilação ou finalidade. Nesta revisão, serão abordados os sistemas de classificação filosóficos e os sistemas de classificação bibliográficos.

Ao se tratar do primeiro tipo, há quem afirme ter sido Platão (~427-347 a.C.) o primeiro a dividir o conhecimento – em quatro bases filosóficas. Outros tomam como pioneiro Aristóteles (~384-322 a.C.) e sua proposta abrangendo as ciências teóricas (matemática, física e teologia), as ciências práticas (moral ou ética, economia e política) e as ciências poéticas (retórica, poética e dialética), inspiradas no pensar, no agir e no produzir, respectivamente.

As premissas de Aristóteles teriam servido de base para Calímaco (~260-240 a.C.) ao desenvolver o catálogo para a biblioteca da Alexandria. “(Calímaco) criou tabelas organizadas alfabeticamente, com nomes de autores e breve biografia. As obras eram classificadas conforme o campo de atuação do autor, como poesia, filosofia, história e outras” (SCHIESSL; SHINTAKU, 2012, p. 74).

Influente nesse assunto foi o filósofo grego Porfírio, que apresentou uma classificação dicotômica, binária, dual. Mais adiante no tempo, as disciplinas estudadas dividiram-se em dois grupos: o *Trivium* (arte ou ciências sensoriais – gramática, dialética e retórica) e o *Quadrivium* (geometria, aritmética, astronomia e música).

Outro autor importante, que viria a influenciar as classificações utilizadas até hoje, foi Francis Bacon (1561-1626), que se baseou na memória (história), na imaginação (poesia) e na razão (filosofia).

Sobre a construção das classes, Piedade reflete:

O homem começou a pensar e a procurar uma explicação para a sua existência, e assim surge a filosofia; incapaz de desvendar o mistério, imaginou a existência de um ser supremo que o havia criado, surge a religião; multiplicando-se o homem passa a viver em sociedade e vêem as ciências sociais; sente necessidade de se comunicar com os companheiros e cria as línguas; passa então a investigar os segredos da natureza e temos as ciências puras; de posse desses conhecimentos procura deles tirar proveito, aparecendo as ciências aplicadas; e, agora, já se sentindo capaz de criar, dá origem às artes e à literatura; finalmente encontramos a história que conta como tudo se passou (PIEADADE, 1983, p. 89).

A classificação tornou-se assunto de interesse de diversos cientistas ao longo dos séculos, como Cassiodoro, Roger Bacon, Leibniz, Hegel, Comte, Wundt até se chegar aos estudos de Cutter, Dewey, Brown, Bliss e Ranganathan, que se dedicaram às classificações bibliográficas.

Com objetivos mais práticos, as classificações bibliográficas foram desenvolvidas para facilitar a ordenação dos documentos nas estantes e nos arquivos, assim como para ordenar referências nas bibliografias e fichas de catálogos. Para Campos (2001, p. 28), os esquemas de classificação bibliográfica teriam dupla função: “a de permitir a organização dos documentos nas estantes e a de representar o conhecimento registrado numa dada área de assunto”.

Os documentos podem ser ordenados segundo os mais variados critérios, como o tamanho ou a cor da encadernação, mas a característica mais útil à classificação de documentos é a ideia apresentada, pois a sua finalidade é facilitar a localização de informações. (...) Quando nos referimos à classificação bibliográfica, subentendemos uma classificação que tem por base os assuntos tratados nos documentos (PIEDADE, 1983, p. 65).

Para Schiessl e Shintaku (2012, p. 59), “a classificação está relacionada diretamente à organização do conhecimento, como forma de categorizar e organizar determinada área, definindo a abrangência do objeto de estudo”. Apesar de a classificação ser usada em diversas disciplinas, é na Ciência da Informação que é tratada como objeto de estudo. Para explicar a função da classificação nos serviços de informação, Schiessl e Shintaku (2012, p. 60) citam Foskett e Bury (1982). Para esses autores a função da classificação nos serviços de informação é ordenar documentos, de modo a facilitar os usuários e serviços. Assim, a classificação é um processo que possibilita a mediação entre os usuários e os documentos, tendo como propósito simplificar a recuperação. Portanto, a classificação apoia a organização física e virtual dos documentos, facilitando a sua recuperação.

No âmbito das áreas relacionadas à Ciência da Informação, os sistemas de classificação ganham relevância e notoriedade especialmente na vida do século XIX para XX. Influenciado pelo trabalho de William Torrey Harris (1835-1909)², o bibliotecário do Amherst College, Melville Louis Kossuth Dewey (1851-1931), desenvolveu um método de classificação que seria publicado anonimamente como *A classification and subject index cataloging and arranging the books and pamphlets of a library* em 1876, mesmo ano em que ajudaria a fundar a *American Library Association* (ALA).

² Bettencourt (2014, p. 25) cita Brown (1898, p. 93) para afirmar que a William Torrey Harris foi creditado o primeiro método de classificação adaptado às necessidades de uma moderna biblioteca, cuja primeira versão foi publicada no *Jornal de Filosofia Especulativa* em 1870.

Melvil Dewey organizou o conhecimento em dez categorias, sendo uma para assuntos gerais e nove para assuntos específicos, como listadas a seguir:

- 000 Generalidades (*General*)
- 100 Filosofia e disciplinas relacionadas (*Philosophy*)
- 200 Religião (*Theology*)
- 300 Ciências sociais (*Sociology*)
- 400 Linguagem ou Filologia (*Philology*)
- 500 Ciências puras (*Natural Science*)
- 600 Tecnologia ou ciências aplicadas (*Useful Arts*)
- 700 Artes, recreação e artes cênicas (*Fine Arts*)
- 800 Literatura (*Literature*)
- 900 História, Biografia, Geografia (*History*)³

A Classificação Decimal de Dewey (CDD) continua em uso, especialmente em bibliotecas públicas.

Em 1891, quando atuava como bibliotecário da Forbes Library, Charles Ammi Cutter (1837-1903) desenvolveu a Classificação Expansiva de Cutter, composta por letras e números. Ela chegou a servir de base para a atualização da classificação da Biblioteca do Congresso Americano, mas perdeu popularidade para a CDD. Na Biblioteconomia, Cutter é mais conhecido pela Tabela de Cutter, relativa à notação de autores.

Diferentes de outros sistemas de classificação, o utilizado pela Biblioteca do Congresso foi desenvolvido a partir da necessidade de organização do acervo, sem a pretensão de classificar todo o conhecimento. A classificação é composta por 21 categorias, enumeradas pelas letras do alfabeto latino. Cada letra representa uma área:

- A – Obras gerais (*General Works*)
- B – Filosofia, psicologia e religião (*Philosophy. Psychology. Religion*)
- C – História e ciências auxiliares (*Auxiliary Sciences of History*)
- D – História Universal (*World History and History of Europe, Asia, Africa etc.*)
- E e F – História das Américas (*History of the Americas*)
- G – Geografia, antropologia e recreação (*Geography. Anthropology. Recreation*)
- H – Ciências Sociais (*Social Sciences*)
- J – Ciências Políticas (*Political Science*)
- K – Legislação (*Law*)
- L – Educação (*Education*)
- M – Música (Music and Books on Music)
- N – Belas artes (*Fine Arts*)
- P – Língua e linguagem (*Language and Literature*)
- Q – Ciências (*Science*)
- R – Medicina (*Medicine*)
- S – Agricultura (*Agriculture*)
- T – Tecnologia (*Technology*)
- U – Ciência Militar (*Military Science*)
- V – Ciência Naval (*Naval Science*)
- Z – Bibliografia, Biblioteconomia e recursos de informação (*Bibliography. Library Science. Information Resources*)⁴

³ *Cataloguing and Arranging the Books and Pamphlets of a Library [Dewey Decimal Classification]*, by Melvil Dewey - <http://www.gutenberg.org/files/12513/12513-h/12513-h.htm>

⁴ Library of Congress Classification Outline - <http://www.loc.gov/catdir/cpsolcco/>

Em 1905, os belgas Paul Otlet (1868-1944) e Henri La Fontaine (1853-1943) lançaram a Classificação Decimal Universal (CDU). Inspirada na CDD e com a anuência de Dewey, a CDU foi adaptada à descrição do Repertório Bibliográfico Universal pretendido pelo Institut International de Bibliographie (IIB). Entre as alterações que sofreu em relação à CDD estão “o corte dos zeros finais, a colocação de um ponto a cada três algarismos e o uso de sinais gráficos e letras”, como descreve Vieira (2014, p. 85).

As classes principais da CDU apresentam-se na seguinte ordem:

- 0 Ciência e Conhecimento. Organização. Informática. Informação. Documentação. Biblioteconomia. Instituições. Publicações
- 1 Filosofia. Psicologia
- 2 Religião. Teologia
- 3 Ciências Sociais
- 4 (vaga)
- 5 Matemática. Ciências Puras
- 6 Ciências Aplicadas. Medicina. Tecnologia
- 7 Arte. Recreação. Entretenimento. Desporto
- 8 Língua. Linguística. Literatura
- 9 Geografia, História e Biografia.⁵

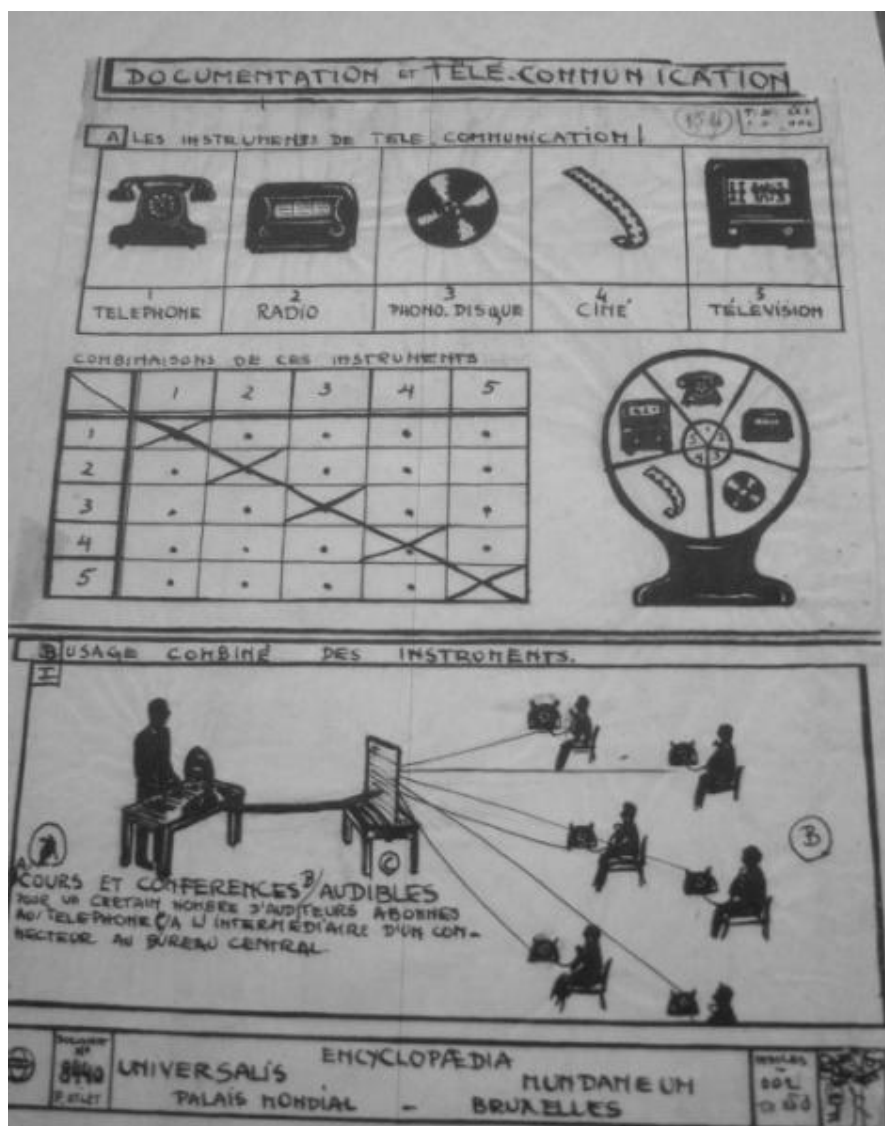
Sobre a criação do IIB, ocorrida em 1895, Pereira escreve:

Otlet e seu companheiro de trabalho – Henri La Fontaine (15 anos mais velho do que ele e agraciado com o Nobel da Paz em 1913) – decidem criar, em Bruxelas, no dia 12 de setembro de 1895, um centro mundial para organização e disseminação do conhecimento internacional, o IIB. Em 1914, o IIB contava com cerca de 700 membros, entre os quais muitos agraciados com o Prêmio Nobel e, em 1924, foi reestruturado, transformando-se em uma federação com cinco membros nacionais – Bélgica, França, Alemanha, Holanda e Suíça. Seu produto mais importante foi uma enorme base de dados chamada *Repertório Bibliográfico Universal* (RBU) (PEREIRA, 2000).

Paul Otlet é uma figura emblemática do campo da documentação. Em uma época ainda desprovida de muitas das tecnologias existentes hoje em dia, Otlet já imaginava uma série de recursos – e de combinações para o uso desses recursos – capazes de ajudar as pessoas a terem acesso à informação. Entre os desenhos guardados no Museu Mundaneum, em Mons, na Bélgica, um chama especialmente a atenção: Otlet já imaginava um aparelho, chamado *Mundaneum*, que combinava diversos recursos (rádio, telefone, televisão, cinema e discos fonográficos), no qual as pessoas poderiam acessar informações registradas (Figura 3).

⁵ The UDC Summary (UDCS) – Disponível em: <<http://www.udcc.org/udcsummary/php/index.php?lang=23>>. Acesso em: 26 jun. 2015

Figura 3 – Desenho de Otlet mostra interação entre documentação e telecomunicação



Fonte: A autora (2012), reprodução de acervo do Museu Mundaneum.

Esboçado em 1895 e desenvolvido parcialmente no início do século XX, o *Mundaneum* abrigaria o repertório mundial do conhecimento em fichas catalográficas, de acordo com a expectativa e a mente profícua de Otlet. Tudo seria desenvolvido no IIB. No projeto de Otlet, os exemplares dos documentos e as microcópias dos livros arrolados no repertório seriam todos depositados no *Mundaneum* (PEREIRA, 1995, p. 2).

Em 1914, o *Répertoire Bibliographique Universel* de Otlet e La Fontaine consistia em mais de 11 milhões de fichas e o IIB tinha aproximadamente 700 membros, como relata Pereira (1995, p. 2) em artigo homenageando os 100 anos do IIB, que a partir de 1938 foi rebatizado como Federação Internacional de Informação e Documentação (FID). Para ordenar esse expressivo repertório, a dupla de pesquisadores desenvolveu a citada CDU.

Atualmente, a CDU é utilizada em bibliotecas e centro de documentação de mais de 130 países em 40 idiomas. Com sua extravagante invenção, como descreve Pereira (1995, p. 3), Otlet pretendia tornar possível a classificação e a recuperação automática de documentos e a manipulação mecânica de todos os dados registrados para obter novas combinações de fatos ou novas relações de ideias.

Um ano após a morte de Otlet e meio século depois de seu *Mundaneum*, outra ideia inovadora chamava a atenção da comunidade científica: o *memex*. Imaginado pelo engenheiro Vannevar Bush, o *memex* foi apresentado no artigo *As we may think*, publicado em julho de 1945 na revista científica *Atlantic Monthly*. No texto, Bush (1945, p. 104) define o *memex* como o “dispositivo no qual o indivíduo armazena todos os seus livros, registros e comunicações, mecanizado e que pode ser consultado com extrema velocidade e flexibilidade”.

Em seu artigo, Bush complementa dizendo que o *memex* é um suplemento ampliado e próximo da memória do indivíduo – anos mais tarde, em 1964, Marshall McLuhan iria desenvolver pensamento semelhante referindo-se aos meios de comunicação como extensões do homem.

Mais do que idealizar um equipamento que solucionaria problemas relacionados à organização e, especialmente, à recuperação da informação, Bush ganhou notoriedade ao questionar os sistemas de indexação alfabéticos e numéricos adotados pelas bibliotecas até então. O pesquisador chamava a atenção para a discrepância que esse modelo tinha em relação à forma como a mente humana funcionava.

Como bem observou Bush, a mente opera por associação e seria esse também o modo como o *memex* trabalharia. Ao adotar esse princípio, afirma Pereira (1995, p. 4), o dispositivo permitiria ao usuário navegar por ele, como se fosse a extensão de sua própria mente. Ao longo dos anos seguintes, o texto visionário serviu de fonte de inspiração para muitos – e continua sendo, mais de meio século depois, referência em diferentes áreas do conhecimento, especialmente nos estudos relacionados à Ciência da Informação. O *memex*, porém, nunca chegou a sair do papel. O princípio de navegar por informações de forma associativa, no entanto, vem sendo empregado pelos mecanismos de busca na “Web”.

Retomando a trajetória das classificações bibliográficas, um sistema diferenciado surgiria nos anos 1930, quando o matemático e bibliotecário indiano Shiyali Ramamrita Ranganathan (1892-1972) desenvolveu e publicou em 1933 a sua Classificação de Dois Pontos – em inglês *Colon Classification*, nome dado em referência ao uso de dois pontos para correlacionar assuntos diferentes. Além do sinal de dois pontos, é utilizada notação mista,

contendo algarismos arábicos, letras maiúsculas e minúsculas, letras gregas e sinais gráficos, somando cerca de 70 caracteres, e indicadores especiais de facetas, conforme descreve Vieira (2014, p. 102).

Até aquele momento no âmbito da Documentação, as tabelas existentes não apresentavam as bases teóricas para sua elaboração. Ranganathan foi o primeiro a evidenciar os princípios utilizados na elaboração de sua tabela, proporcionando uma verdadeira revolução na área da classificação bibliográfica. Na verdade, ele não elabora somente um trabalho teórico para explicar a construção da tabela, mas apresenta uma teoria sólida e fundamentada para dar à classificação bibliográfica um *status* que a eleva a disciplina independente (CAMPOS, 2001, p. 27).

“A principal característica desse esquema são as facetas e focos para a representação dos assuntos de um documento. Facetas são os aspectos em que o assunto pode ser contextualizado, como personalidade, matéria ou propriedade, energia, espaço e tempo.” (SCHIESSL; SHINTAKU, 2012, p. 70) Conhecidas como PMEST, as categorias instituídas por Ranganathan referem-se a:

Personalidade (*personality*) – facetas de um assunto que representa a sua essência, por exemplo, biblioteca (classe principal) e biblioteca universitária (personalidade).
 Matéria (*matter*) – matérias e substâncias que constituem as coisas, por exemplo: livros, periódicos.
 Energia (*energy*) – processos, operações, ações, técnicas, métodos, fenômenos, por exemplo: processamento técnico, empréstimo.
 Espaço (*space*) – divisão geográfica.
 Tempo (*time*) – divisão cronológica.
 (RANGANATHAN, 1963).

O modelo estabelecido por Ranganathan dá início aos sistemas modernos de classificação.

Como descreve Campos (2001, p. 30), os classificacionistas anteriores a Ranganathan costumavam organizar seus esquemas de classificação a partir de assuntos representativos da literatura naquele momento histórico, ou seja, o que mais aparecia nos livros existentes em determinado período, o conhecimento já estabelecido. Quando surgiram assuntos novos, tornava-se difícil encaixá-los nas classificações. Este foi um dos problemas identificados por Ranganathan durante seus estudos de Biblioteconomia. “Pode-se afirmar que não ocorria ligação entre o conhecimento e as classificações naqueles esquemas, mas entre os assuntos dos documentos e as classificações” (CAMPOS, 2001, p. 30).

2.2.1.2 Vocabulários controlados e tesouros

Tesouro, armazenamento, repositório (VIEIRA, 2014, p. 143), haveres, local onde se acumulam bens, materiais ou não (SCHIESSL; SHINTAKU, 2012, p. 88). Tendo como origem o termo *thesaurós* no grego que levou a *thesaurus* no latim, o tesouro agrupa termos de acordo com as ideias que estes exprimem, diferente de dicionários organizados em ordem

alfabética. Um tesauro é concebido como um instrumento prático para auxiliar na indexação e na recuperação de documentos, assim como o são os cabeçalhos de assuntos e os vocabulários controlados.

Vocabulários controlados, porém, são simples listas de palavras-chave, sem nenhum tipo de relação e um controle mínimo dessas palavras. São listas de termos equivalentes, na forma de um círculo de sinônimos – um conjunto de palavras equivalentes no caso da recuperação de informação, pois nem sempre as palavras são totalmente parecidas, mas dentro de determinado contexto podem ser vistas como tal.

Cabeçalhos de assuntos também são listas de termos, como descrevem Colepícolo et al. (2006, p. 1), mas com maior controle sobre os termos e agregando relações diretas entre estes. Os cabeçalhos de assunto são empregados para designar os termos que resultam de uma escolha e de uma elaboração da linguagem posterior a sua extração do texto dos documentos, como os apresentam Guinchat e Menou (1994, p. 145).

Cesarino e Pinto (1978, p. 274) listam fatores que determinaram o aparecimento e a formação de cabeçalhos de assunto: títulos de obras que não representavam de forma adequada o assunto das mesmas; problemas relativos às subdivisões de assuntos; obras com mais de um assunto; livros de assuntos relacionados; obras que relacionavam os assuntos a lugares e épocas diferentes.

As listas podem conter tanto termos escolhidos quanto sinônimos de linguagem natural junto aos termos correspondentes. “Quanto mais organizado for o vocabulário, maior é a possibilidade de controlá-lo, isto é, de evitar os duplos empregos, a multiplicação de palavras muito precisas que dizem respeito apenas a um número reduzido de documentos” (GUINCHAT; MENO, 1994, p. 145).

O que diferencia tesauro de cabeçalho de assunto são os tipos de relações existentes entre os termos. Um cabeçalho de assunto apresenta somente relações hierárquicas entre termos, enquanto no tesauro, além das relações hierárquicas, também são encontradas relações de equivalência e de associação (COLEPÍCOLO et al., 2006).

Para Colepícolo et al. (2006, p. 1), um tesauro se define como um vocabulário controlado, que representa hierarquias, relações de equivalência, pertinência e associações entre os termos, com objetivo de auxiliar o usuário potencial a encontrar a informação de que necessita com a menor margem de erro possível.

Um tesauro distingue-se de um simples vocabulário por duas características. A primeira é que as palavras nele listadas não descrevem, mas antes significam. Cada termo nomeia um conceito. Sendo assim, esses termos que designam conceitos não são simples palavras, tornam-se “termos” ou “descritores”. A segunda característica é que todos estão relacionados entre si, nenhum termo pode figurar no tesauro sem estar relacionado a algum outro, sendo essa relação determinada pela definição de seu conceito (MIRANDA, 2005, p. 147).

Cavalcanti (1978, p. 27), citada por Vieira (2014, p. 143) e por Schiessl e Shintaku (2012, p. 91), define tesauro como uma lista estruturada de termos relacionados, utilizada por profissionais da informação para descrever um documento com a especificidade desejada e permitir aos usuários a recuperação da informação pretendida. Segundo Guinchat e Menou (1994, p. 146), “tesauro é um conjunto controlado de termos entre os quais foram estabelecidas relações hierárquicas e relações de vizinhança. Também são estabelecidas, quando necessário, relações de sinonímia com palavras da linguagem natural”.

As relações hierárquicas de um tesauro são relações de ordenação entre os termos, ou seja, de superordenação (acima de), subordinação (abaixo de) e coordenação (na mesma ordem, igual a). Em um jornal, por exemplo (Figura 4), a redação é superordenada em relação às editorias. Já estas são subordinadas à redação, mas encontram-se no mesmo nível, ou seja, em nível coordenado, em relação às demais editorias do jornal.

Figura 4 – Relações hierárquicas em um tesauro



Fonte: A autora (2015), a partir do modelo de Colepícolo et al. (2006, p. 2).

De acordo com Colepícolo et al. (2006, p. 2), as relações de equivalência envolvem o estudo e delimitação de termos diferentes com um mesmo significado e termos idênticos com significados diferentes, entre outras relações de equivalência entre termos já estabelecidos pela gramática das línguas, ou seja, sinônimos, antônimos, parônimos e homônimos.

Para a NISO Z39.19:2005,

Tesauro é um vocabulário controlado, organizado em ordem e estruturas conhecidas de tal maneira que os relacionamentos de equivalência, homografia, hierarquia e associações entre os termos sejam dispostos claramente e representados por indicadores de relacionamentos padronizados e recíprocos.

Schiessl e Shintaku apresentam uma síntese dos tipos de relacionamentos que podem ser estabelecidos e mostram que os relacionamentos são mútuos:

- Tesouro é uma ferramenta voltada a organização e recuperação da informação.
- Atende aos usuários: ao indexador que estabelece e define os termos e seus relacionamentos que representam documentos e ao consulente, que busca informação auxiliado pela ferramenta.
- É um tipo de vocabulário controlado expresso em linguagem documentária que formaliza os termos simples ou compostos e os relacionamentos entre eles.
- A estrutura dos tesouros permite a organização lógica e semântica de um domínio.
- A mesma estrutura também permite que os significados dos termos sejam limitados de forma unívoca.
- Essa restrição aos significados dos termos visa neutralizar a sinonímia, a polissemia e outras características comuns à linguagem natural.
- Tesouro oferece uma representação da estrutura conceitual de um domínio do conhecimento (SCHIESSL; SHINTAKU, 2012, p. 92).

Currás (1995, p. 77) relembra que em tempos não muito remotos, mais especificamente entre as duas grandes guerras mundiais no século XX, o mais importante eram os documentos que continham a informação, não a informação que estava contida nos documentos. Assim, a classificação era uma das principais tarefas dos profissionais que trabalhavam com tais documentos. A autora relata que já naquela época tinha-se início o pensamento de que os leitores/usuários queriam, no fundo, saber dos conteúdos dos livros e não sobre o suporte que carrega a informação.

De acordo com Currás, esta mudança de mentalidade ganhou mais forças após a II Guerra Mundial, quando se retomou o progresso tecnológico.

Verdadeiramente, as classificações por assunto existentes tornaram-se demasiadamente rígidas e excessivamente estáticas para enquadrar nelas as novas invenções e descobertas, os novos conceitos, os novos termos que apareciam diariamente. Assim surgiu a ideia de tirar do documento seu conteúdo e representá-lo por palavras (palavras-chave), termos, que seriam submetidos a uma ordenação posterior (CURRÁS, 1995, p. 77).

A partir da aplicação desta ideia, o paradigma passava de uma linguagem natural, a dos documentos, para uma linguagem estruturada, a da informação neles contidas. Nesta época, de acordo com Currás, aparecerem expressões como “linguagem controlada”, “linguagem normalizada”, “linguagem documentária” (usada para fins classificatórios).

Para resolver problemas característicos do uso da linguagem natural – mapeando, por exemplo, os termos que representem o mesmo conceito, selecionando um termo apenas como padrão e os restantes como sinônimos, além de estabelecer relações entre esses termos e outros a estes relacionados – é que os tesouros se apresentam como excelente alternativa.

Schiessl e Shintaku (2012, p. 94) citam duas funções dos tesouros. A primeira é a capacidade de representar os assuntos do documento, ou seja, o indexador analisa o documento, identifica os termos que melhor representam o assunto e os exprime na linguagem específica do tesouro. A segunda é a competência de também representar as buscas realizadas por usuários, relacionados aos assuntos. “Essa função proporciona um refinamento às buscas,

pois mesmo que o usuário não saiba um termo técnico utilizado no tesouro, a estrutura de relacionamento entre os termos possibilita essa conversão e, assim, recupera a informação desejada” (SCHIESSL; SHINTAKU, 2012, p. 94).

As grandes vantagens dos tesouros são sua especificidade, maleabilidade e capacidade de descrever as informações de forma completa. Mas sua elaboração requer um trabalho considerável. Raramente um tesouro pode ser reutilizado sem sofrer modificações. Atualmente, existe uma grande proliferação deste tipo de linguagem, o que dificulta a comunicação entre unidades de informação e exige grandes esforços de normalização, de pesquisa e de compatibilidade entre estes instrumentos (GUINCHAT; MENO, 1994, p. 147).

Um dos maiores desafios na criação de tesouros é a seleção de termos. De acordo com Guinchat e Menou, palavras e frases extraídas do domínio de conhecimento determinam a formulação de termos. Um dos métodos recomenda coletar na literatura da área a ser tratada, pois o contexto é imprescindível. Outra fonte pode ser a tabela de classificação. Além disso, não se pode deixar de envolver os usuários no processo de criação de um tesouro.

Para Carlan e Medeiros (2011, p. 56), os tesouros são estruturas consolidadas e padronizadas por normas internacionais, que tratam da construção, formatação e manutenção de tesouros monolíngues e multilíngues. Tendo como ponto de partida essas normas, as autoras sintetizam a estrutura de tesouros em duas partes: a base teórica e a base técnico-operacional.

Base teórica é a essência e a fundamentação para a construção de tesouros: conceito, termo, categorias e faceta. Base técnico-operacional é abordagem mais prática no desenvolvimento de tesouros: planejamento, coleta de termo, controle terminológico, estabelecimento de relações entre conceitos e formas de divulgação e publicação.

2.2.1.3 Códigos de catalogação

Estima-se que um dos primeiros catálogos tenha sido organizado por Calímaco, bibliotecário da Biblioteca de Alexandria, por volta de 250 a.C. Há dúvidas, porém, se o resultado do trabalho de Calímaco era um catálogo ou apenas uma bibliografia da literatura grega. Como apontam Mey e Silveira (1987, p. 44), “os gregos foram responsáveis pela introdução do conceito de autor de uma obra como ponto de acesso a ela, princípio que permanece até nossos dias”.

Por séculos, a catalogação não teve grandes avanços – sendo designada, na prática, como listas de obras ou inventário de coleções. Foi com o desenvolvimento da imprensa – ocorrido após a criação dos tipos móveis pelo alemão Johannes Gutenberg (1398-1468) por volta de 1450 –, que surgiu a necessidade de organizar os livros de maneira mais cuidadosa.

De acordo com a pesquisa de Fiuza (1987, p. 45), precursor nessa tarefa foi o suíço Conrad Gessner (1516-1565), que, em 1545, publicou uma bibliografia arranjada alfabeticamente por autor, à qual foi acrescentado um índice de assunto em 1548.

Mey e Silveira (2009, p. 67) destacam os papéis do monge beneditino alemão Florian Trefler, que publicou um tratado sobre a manutenção de uma biblioteca em 1560, e do livreiro inglês Andrew Maunsell (1560-1604), que estabeleceu elementos básicos para a descrição bibliográfica em 1595. Fiuza destaca ainda o bielorusso Seymour Lubetzki (1898-2003), que descreveu a ideologia da catalogação desde o final do século XVII até meados do século XIX.

A biblioteca é uma fonte de livros e o catálogo é um auxiliar para localizar um livro dentro da biblioteca. Um livro é normalmente identificado pelo nome do autor – real ou assumido – e pelo título. Portanto, raciocinou-se, entradas sob o nome do autor e sob o título encontrado na folha de rosto ou em outra parte do livro, em conjunto com um símbolo para indicar sua localização nas estantes, permitiriam encontrar rapidamente um livro (LUBETZKY, 1979 apud FIUZA, 1987, p. 45).

No século XIX, Antonio Panizzi (1797-1879) teve um papel central no avanço da catalogação ao ser contratado como bibliotecário assistente pelo Museu Britânico em 1831. Panizzi já havia trabalhado no catálogo da *Royal Society of London*. Após quatro anos, ele lançava o novo catálogo do Museu Britânico, no qual incluía “suas famosas 91 regras, onde estava contida sua ideologia sobre a catalogação”, como ressalta Fiuza (1987, p. 46). Panizzi defendia as conexões entre as obras, não a representação dos livros como entidades separadas das demais.

No cenário norte-americano, Charles Coffin Jewett (1816-1868) e Charles Ami Cutter protagonizaram o desenvolvimento inicial da área no século XIX. Jewett adequou as ideias de Panizzi para o contexto em que vivia, idealizou um sistema de fichas por meio de placas estereotipadas de cobre e foi defensor de um catálogo coletivo nacional.

Cutter estabeleceu um código racional e completo com 368 regras para um catálogo dicionário impresso (*Rules for a Dictionary Catalogue*), a partir das quais o usuário poderia encontrar no acervo um livro de que já conhecia o autor, o título ou o assunto, ver o que a biblioteca tinha sobre determinado autor, assunto ou espécie de literatura e obter ajuda para escolher um livro, fosse pela edição ou pelo caráter. A obra serviu como base para os códigos de catalogação utilizados até hoje.

Em 1901, como descrevem Mey e Silveira (2009, p. 73), um fato mudou a perspectiva dos códigos: “a impressão e venda de fichas pela Biblioteca do Congresso”. Ao invés de cada biblioteca confeccionar suas fichas, elas compravam as fichas padronizadas, que seguiam as regras estabelecidas por Panizzi, Cutter e Jewett e pelo *LC: Cataloging rules: author and title entries*.

Outros países lançaram suas regras e seus códigos. No Brasil dos anos 1940, até os anos 1980, teve grande influência o Código da Vaticana. Este era o código seguido no curso de Biblioteconomia da Biblioteca Nacional. Já o ensino da Biblioteconomia em São Paulo tinha influência norte-americana, sendo utilizado o código da ALA.

Houve tentativas no país de criar um código próprio, mas a ideia não prosperou. Mey e Silveira (2009, p. 75) afirmam que apenas normas de cabeçalhos de nomes em língua portuguesa foram obtidas e publicadas em apêndice às duas edições brasileiras do código anglo-americano. Como apontado por Mey e Silveira, marcos importantes na trajetória da catalogação foram:

a) Em 1927, a fundação da Federação Internacional de Associações Bibliotecárias (IFLA, *International Federation of Library Associations*) – renomeada em 1976 Federação Internacional de Associações e Instituições Bibliotecárias (*International Federation of Library Associations and Institutions*). A sigla não sofreu alteração.

b) Em 1946, a fundação da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (Unesco), que tinha entre suas funções manter, desenvolver e disseminar o conhecimento, por meio da cooperação internacional, para o acesso a todos os materiais publicados. A Unesco criou o Controle Bibliográfico Universal, gerenciado pela IFLA.

c) Em 1954, a criação do Instituto Brasileiro de Bibliografia e Documentação (IBBD), que resultou na criação de um curso de especialização em Curso de Especialização em Informação e Documentação (CDC), na difusão do Código da Vaticana, no Serviço de Intercâmbio de Catalogação (SIC), entre outras atividades. Em 1975, o IBBBD tornou-se Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), tendo entre suas funções o comando da Biblioteca Digital de Teses e Dissertações (BDTD) e do Catálogo Coletivo Nacional de Publicações Seriadas, além da atribuição do ISSN (número internacional normalizado do periódico).

Como reportam Mey e Silveira (2009, p. 77), a década de 1960 trouxe o desenvolvimento dos recursos computacionais, inserindo mudanças significativas no processo de catalogação. A Biblioteca do Congresso passou a desenvolver o Projeto MARC (*Machine Readable Cataloging*, 1965-1966) e MARCII (1968), este junto com a Biblioteca Britânica. Como ressaltam, “o MARC não é um tipo de catálogo, nem um método de catalogação, mas um formato, um padrão de entrada e manuseio de informações bibliográficas em computador”. De acordo com Vieira (2014, p. 113), o MARC fornece o mecanismo pelo qual os computadores interpretam, utilizam e trocam entre si informações bibliográficas, e que juntos formam a base da maioria dos catálogos de bibliotecas usados atualmente.

Outro padrão criado nesta época e em uso até hoje é o Código de catalogação anglo-americano (AACR), elaborado pela ALA, pela *Canadian Library Association* e pela *Library Association* (Inglaterra). Atualmente, a maioria das bibliotecas no Brasil utiliza o AACR2, tradução publicada em 2004 a partir da edição de 2002. Outros códigos já existem, mas ainda são pouco utilizados nas unidades de informação brasileiras, como o RDA (*Resource Description and Access*, Descrição e acesso aos recursos) e o Código internacional de catalogação (*International Cataloguing Code*).

Ainda nos anos 1960, como descrito por Mey e Silveira (2009, p. 79), é apresentada a Descrição Bibliográfica Internacional Normalizada (ISBD, *International Standard Bibliographic Description*), que padroniza as informações contidas na descrição bibliográfica a fim de tornar seu reconhecimento possível por computadores. A versão mais recente data de julho de 2011.

Entre as novidades desenvolvidas nas últimas décadas na área, tornou-se um marco na catalogação contemporânea a publicação da FRBR (*Functional Requirements for Bibliographic Records: final report*, Requisitos Funcionais para Registros Bibliográficos), relatório final da 63ª Conferência Geral da IFLA, em 1998. Os FRBR influenciaram as revisões das ISBDs e dos códigos de catalogação, a adoção de nova terminologia na área e o desenvolvimento de novos modelos conceituais para o universo bibliográfico.

2.2.1.4 Ontologias

O termo ontologia deriva do grego *ontos* (ser) e *logos* (palavra, discurso, razão). Ao tratar sobre ontologia é comum na Ciência da Informação considerar primeiramente a definição usada na filosofia, ou seja, ontologia como estudo dos entes, das coisas como verdadeiramente são, do “ser” ou “o que existe” na natureza, como escrevem Schiessl e Bräscher, baseados em Chauí (2003), para logo em seguida questionarem:

Mas, afinal, o que é ontologia para a Ciência da Informação? Numa proposta, Gruber (1993) afirma que ontologia é a especificação formal e explícita de conceitualização. Borst (1997) complementa que essa conceitualização deve ser compartilhada, na qual os grupos de objetos e suas relações são refletidos por vocabulário representacional. (...) Daconta et al. (2003) asseguram que ontologia define termos e conceitos comuns na representação e descrição de determinada área de conhecimento e, assim, padroniza o significado (SCHIESSL; BRÄSCHER, 2011, p. 301).

Além de definir ontologia como a especificação de uma conceitualização, Gruber (1993, p. 1) afirma ainda que uma ontologia é a descrição dos conceitos e das relações que podem existir para um agente ou uma comunidade de agentes.

Tristão, Fachin e Alarcon esmiúçam a definição de Gruber:

[...] conceitualização refere-se a um modelo de fenômeno abstrato no mundo por ter identificado os conceitos relevantes daquele fenômeno; explícito significa que o tipo dos conceitos usados e as restrições no seu uso são definidos explicitamente; formal se refere ao fato de que a ontologia deveria ser lida pela máquina; e compartilhado reflete que ontologia deveria capturar conhecimento consensual aceito pelas comunidades. (TRISTÃO; FACHIN; ALARCON, 2004, p. 162)

Gruber ressalta que mais importante que a definição é entender para que serve uma ontologia, por qual motivo foi criada. “Meus colegas e eu temos criado ontologias com o propósito de possibilitar compartilhamento de conhecimento e reuso” (GRUBER, 1993, p. 1).

Para Carlan (2010, p. 52), ontologias apresentam-se como modelos de relacionamento de entidades em um domínio particular do conhecimento, sendo objetivo de sua construção suprir a necessidade de um vocabulário compartilhado pelo qual as informações possam ser trocadas e também reutilizadas pelos usuários de uma comunidade. Carlan afirma ainda que as ontologias promovem e facilitam a interoperabilidade entre sistemas de informação, ou seja, é possível compartilhar e reutilizar o conhecimento entre sistemas.

Pode-se descrever ontologia pela definição de um grupo de termos representativos de determinado domínio, sendo que a tarefa desse domínio é simbolizar formalmente essa existência. Assim, uma ontologia reflete uma visão de mundo, uma maneira limitada, porém útil, de certo domínio. Por representarem um domínio determinado, as ontologias não seguem sempre a mesma estrutura. Porém, costumam ter características e componentes básicos comuns, como ressaltam Almeida e Bax (2003, p. 9). Tais componentes são classes, relações, axiomas e instâncias.

Dentro da Ciência da Informação, como avalia Campos (2010, p. 221), ontologia determina um vocabulário comum para uma comunidade que precisa compartilhar informação em determinado domínio, além de incluir definições de conceitos básicos de tal domínio e as relações entre estes de forma a tornarem-se interpretáveis pela máquina. A autora ressalta que as descrições possibilitam o entendimento do significado de um termo por um agente inteligente e que assim sejam estabelecidas inferências sobre os significados, “pois a definição é composta de características de conceitos, que são também conceitos que se relacionam formando o entendimento semântico dos termos em questão” (CAMPOS, 2010, p. 222).

A ontologia é alicerçada pela representação das características fundamentais do termo, que são utilizados para descrever e representar uma determinada área do conhecimento. Uma ontologia pode ser usada por pessoas, bancos de dados, em técnicas e aplicações de raciocínio indutivo e inferências que necessitam compartilhar informações dentro de um domínio, observam Carlan e Medeiros (2011, p. 57).

Carlan e Medeiros (p. 58) resgatam Noy e McGuinness (2001) para apresentar cinco motivos pelos quais o desenvolvimento de ontologias torna-se importante e considerável: a) compartilhamento de conhecimento comum em estruturas de informação entre outros povos ou para os agentes de *software*; b) reuso do conhecimento; c) realização de inferências em um domínio de conhecimento; d) separação entre o conhecimento de domínio e o conhecimento operacional; e e) análise do conhecimento estruturado tendo como resultado respostas mais relevantes.

2.3 ORGANIZAÇÃO DA INFORMAÇÃO

Profissionais da informação, como cientistas da informação e bibliotecários, envolvem-se há décadas no desenvolvimento de sistemas de organização do conhecimento. Nas páginas anteriores, recapitulou-se algumas das soluções desenvolvidas para melhor representação do conhecimento.

É a partir dos resultados da atuação desses profissionais que se torna possível a aplicação, o viés mais prático, ou seja, a representação da informação, parte essencial da organização da informação. É nela que se encontra o processamento técnico de documentos, o tratamento da informação, seja por sua forma (representação descritiva = catalogação) ou conteúdo (representação temática = indexação, classificação e catalogação de assuntos).

Esta etapa, como apontam Fujita, Rubi, Boccato (2009, p. 19), também pode ser chamada de tratamento documentário, inserida em um conjunto de operações denominado ciclo documentário (ou cadeia documental), composto por coleta de documentos, identificação, gravação, organização, armazenamento, recuperação, conversão intelectual dos materiais e, por fim, difusão da informação.

Fujita, Rubi, Boccato (2009, p. 20) resumem o ciclo documentário em três grandes operações: coleta (localização, seleção e aquisição de documentos); tratamento (processamento dos documentos); e difusão (empréstimo, levantamentos bibliográficos, comutação etc.). Ao se tomar o cenário das hemerotecas digitais criadas por Bibliotecas Nacionais, imagina-se que o processo poderia ser da seguinte forma: coleta (depósito legal), tratamento (seleção, preparação dos originais, microfilmagem/digitalização, catalogação e indexação) e difusão (disponibilização *on-line* do material).

A etapa de tratamento da informação será abordada nos subcapítulos seguintes, em especial a catalogação e a indexação/catalogação de assuntos.

2.3.1 Catalogação

A catalogação é processo importante na OI, pois individualiza um item informacional, um documento, uma informação como coisa. É feita com base na descrição das características dos documentos, sendo uma das tarefas centrais dos bibliotecários. A representação descritiva realizada na catalogação torna o documento único em um acervo. Desta forma, o usuário tem maior facilidade em localizá-lo em uma coleção.

Quando se cataloga um documento, identificam-se autor e título, elabora-se resumo e selecionam-se palavras-chave, entre outros procedimentos, estamos fazendo a representação da informação que consta no documento. A partir do contato com essa representação, o usuário poderá inferir, por um processo de indução, sobre os elementos que constituem o documento representado. A representação não substitui o documento, mas possibilita identificar seus atributos fundamentais, os quais orientam o usuário para uma tomada de decisão sobre a necessidade de consulta do documento que originou a representação (LIMA; ALVARES, 2012, p. 36).

O profissional responsável pela catalogação deve ser um perito na leitura técnica do documento, ou seja, deve ter a habilidade de extrair do documento – seja este, livro, CD, DVD ou publicação periódica – suas características (autor, título, editora, ano de publicação, volume, formato, número de páginas, título original, tradutor etc.), assim como fazer uma descrição minuciosa sobre o assunto tratado.

Como apresentado no *Dicionário de Biblioteconomia e Arquivologia* (CUNHA; CAVALCANTI, 2008, p. 70), no passado “a catalogação tratava-se de um processo técnico para registro e descrição de itens tendo em vista a organização de catálogos”. Até o início do uso de computadores, tal descrição era apresentada em forma de metadados em fichas catalográficas.

No âmbito da catalogação, como descreve Herrero Pascual (1996, p. 66-73), jornais e revistas recebem tratamento semelhante aos demais itens do acervo. A descrição bibliográfica engloba título e menção de responsabilidade, edição, data de publicação, data de início de publicação, local de publicação, descrição física (número de páginas, tamanho, se contém ilustrações etc.), série, notas obrigatórias (periodicidade, tradução, variações do título, se houve troca de título, se foi interrompida), notas opcionais (editada em várias línguas, edições suplementares etc.) e número normalizado (ISSN).

O ponto de acesso principal de um periódico costuma ser o título, mas podem ocorrer entradas por instituição ou por autores (quando a publicação apresenta textos de apenas um autor, por exemplo, mas isto é menos frequente em jornais).

A catalogação de jornais e revistas mostra-se como o processo mais importante para a recuperação da informação contida nesse tipo de item, pois em praticamente todos os casos a busca é realizada a partir do título, da data e do local da publicação.

2.3.2 Indexação e catalogação de assunto

A indexação identifica e descreve um documento a partir dos assuntos nele apresentados. Como afirmam Cardoso e Santos (2012, p. 186), a indexação agrega valor à informação, organizando-a e tornando-a acessível. Indexação é a prática de representar um documento por meio de termos baseados em conceitos bem definidos. Indexar é organizar a informação, traduzindo o que está no documento para uma linguagem acessível e padronizada, facilitando assim a busca e recuperação da informação pelo usuário que dela necessita.

Para Miranda (2005, p. 55), tratamento temático da informação ou indexação é uma operação que consiste em identificar sobre o que trata o conteúdo de um documento. Para assim obter a síntese, mediante a análise e a representação de conceitos, palavras-chave, termos, por meio de uma linguagem de indexação visando a localização e recuperação de documentos/informação em quaisquer ambientes.

Assim como em outros instrumentos desenvolvidos e utilizados por cientistas da informação, o conceito é a matéria-prima da indexação, sendo representado cada conceito por um termo. “A atribuição de termos pelo indexador se refletirá na recuperação da informação” (CARDOSO; SANTOS, 2012, p. 188).

Rowley (2002, p. 162) define indexação como o processo de atribuir termos ou códigos de indexação a um registro ou documento, termos ou códigos esses que serão úteis posteriormente na recuperação do documento ou registro. Condicionar a recuperação da informação a satisfazer as necessidades informacionais dos usuários é a finalidade da indexação, segundo Rubi (2009, p. 97).

Para Cesarino e Pinto,

[...] o objetivo da indexação é, pois, em essência, a transformação de uma expressão da linguagem natural em uma linguagem de indexação. Parte de um texto compreendido por palavras, frases, parágrafos em sequência linear e termina numa condensação do todo em linguagem documentária. A eficiência dos sistemas de recuperação de informações depende muito da estrutura e composição dessa linguagem. Se ela for composta de classes muito abrangentes e gerais, o sistema terá um grau muito superficial de especificidade, levando conseqüentemente a uma alta revocação e baixa precisão. Já se a linguagem se compuser de um número muito grande de classes, estas serão muito específicas e o sistema terá um maior grau de precisão e conseqüentemente baixará o grau de revocação (CESARINO; PINTO, 1978, p. 272).

O processo de indexação inclui a análise do item a ser indexado, ou seja, a leitura do texto para identificação dos conceitos, a síntese, relacionada à elaboração de resumos quando é o caso, e a representação por meio do uso de linguagens documentárias e baseada em políticas de indexação capazes de nortear o trabalho do bibliotecário com diretrizes e critérios

bem definidos. Tais políticas vão determinar o grau de especificidade, exaustividade e, conseqüentemente, a capacidade de revocação e de precisão na recuperação da informação.

Rubi (2009, p. 85) esclarece que a exaustividade diz respeito ao número de termos atribuídos como descritores do assunto do documento, ou seja, em que medida todos os assuntos discutidos no documento são reconhecidos durante a indexação e traduzidos na linguagem documentária da biblioteca. Quanto mais exaustiva for a indexação, mais termos ela vai empregar. Já a especificidade relaciona-se ao nível de abrangência que a biblioteca e a linguagem documentária permitem esmiuçar os conceitos identificados no documento.

A indexação realizada de maneira mais específica resultará, portanto, em uma recuperação com níveis de revocação menores e com um índice maior de precisão, ou seja, mesmo sendo um número reduzido de documentos, são exatamente estes que correspondem às questões de busca do usuário (RUBI, 2009, p. 86).

Para Rowley (2002, p. 169), a linguagem de indexação é composta por termos ou códigos usados como pontos de acesso em um índice, e a linguagem de busca, por termos usados por um consulente ao especificar uma necessidade de busca.

A linguagem de busca do usuário deve ser compatível com a linguagem documentária do sistema, e esta deve representar as necessidades de informação do usuário, construídas por seus modelos mentais influenciados por seu meio. Sob essa perspectiva, vemos a linguagem documentária como um canal de comunicação social, imbuída de valores, em que os conceitos representados por termos devem refletir a cultura do indivíduo e do ambiente em que ele está inserido e da área de conhecimento a que ela corresponde (BOCCATO, 2009, p. 123).

Para que a recuperação da informação seja plena, é necessário que as duas linguagens estejam alinhadas. Ao abordar a questão das linguagens de indexação, a autora lista vantagens e desvantagens (Quadro 1) em uma comparação das linguagens de indexação não-controladas e do vocabulário controlado:

Quadro 1 – Comparação entre linguagens de indexação e vocabulário controlado

	Vantagens	Desvantagens
Linguagens de indexação não-controladas	Baixo custo; processo de buscas simplificado; possível fazer buscas no conteúdo total da base de dados; toda palavra tem valor de recuperação igual; sem erros humanos de indexação; sem demora na incorporação de termos novos.	Maior carga de trabalho para o indexador; podem se perder informações que estejam incluídas de forma implícita, mas não explícita no texto; ausência de vínculos do genérico para o específico; é preciso conhecer o vocabulário da disciplina.
Vocabulário controlado	Resolve muitos problemas semânticos; permite que relações de gênero-espécie sejam identificadas; mapeia áreas do conhecimento.	Custo alto; possíveis inadequações de cobertura; erro humano; possibilidade de vocabulário desatualizado; dificuldade de incorporar todas as relações relevantes entre termos.

Fonte: Rowley (2002, p. 171).

Miranda afirma que o processo de elaboração de uma estratégia de busca de informação é similar ao processo de indexação.

Primeiro, o profissional da informação analisa a necessidade de informação do usuário, que é decomposta em aspectos dos assuntos desejados, interpretando o que realmente o usuário está buscando. Em um segundo momento, traduz o resultado de sua análise para a linguagem do sistema de recuperação da informação (SRI) em termos apropriados para a busca da informação. Em um ambiente atual decidirá quais descritores ou cabeçalhos de assunto irá utilizar para recuperar a informação por meio de catálogos e/ou índices de assunto. Em ambientes eletrônicos e digitais decidirá sobre quais termos da linguagem do sistema utilizará para obter elementos que possam satisfazer às necessidades de informação do usuário, e nos ambientes virtuais utilizará a linguagem natural (MIRANDA, 2005. p. 66).

Para que a indexação realmente facilite a interação do usuário com o acervo, o profissional da informação que a realiza precisa ficar atento ao público-alvo de sua unidade de informação. As informações documentárias são extraídas após uma análise criteriosa do documento. Os termos então serão escolhidos ou no próprio documento ou de um vocabulário controlado – sendo o uso deste último muito mais apropriado e eficaz.

Como relembra Cunha (2008, p. 7), o paradigma da unidade representativa de informação era o livro e não os seus capítulos. “Entretanto, em uma coleção digital heterogênea, os níveis de representação do conteúdo alcançam detalhes inimagináveis, podendo ser um mapa, uma figura, um filme, um slide, um capítulo e mesmo um verbete de uma obra de referência.” Os níveis de representação da informação a serem adotados serão determinados pela política de indexação da unidade de informação.

Lancaster aponta fatores que influenciam na qualidade da indexação:

[...] ligados ao indexador (conhecimento do assunto e do conteúdo temático tratado, entendimento da terminologia abordada e das necessidades do usuário, experiência, concentração, capacidade e compreensão de leitura); ao documento (conteúdo temático, complexidade, língua e linguagem, expressão, apresentação e sumarização); ao vocabulário (especificidade, ambiguidade, qualidade de estrutura, qualidade do vocabulário, disponibilidade de instrumentos auxiliares); ao processo (tipo de indexação, regras e instruções, produtividade exigida, exaustividade da indexação); ambientais (refrigeração, iluminação, ruído) (LANCASTER, 2004, p. 81).

Neves (2012, p. 40) ressalta que a indexação é um processo subjetivo que depende, em grande parte, do desempenho do indexador. Tal profissional é instruído durante a sua formação sobre procedimentos que deverá desempenhar a fim de identificar o conceito ou conceitos a partir da leitura técnica. Os conceitos são encontrados em determinadas partes do documento consideradas importantes, garantindo que informação alguma seja negligenciada. As partes em questão compreendem título e subtítulo, resumo, sumário, introdução, ilustrações, diagramas, tabelas e seus títulos explicativos, palavras em destaque e as referências do documento.

A leitura do documento é apenas a primeira parte do processo de indexação, como afirma Neves. O processo completo é composto por:

[...] compreensão do texto (na qual a leitura se insere), identificação dos conceitos e escolha dos conceitos que representarão o documento, geração de texto (tradução do conteúdo identificado para a linguagem de indexação e/ou sua expressão na forma de um resumo) e representação do conteúdo (criação de linguagens, como os tesouros, sistemas de classificação, dentre outros) (NEVES, 2012, p. 40).

Itens semelhantes aos apontados na Norma Brasileira 12.676 da Associação Brasileira de Normas Técnicas (ABNT), a qual determina o método para análise de documentos e engloba três estágios:

- Exame do documento e estabelecimento do assunto de seu conteúdo.
- Identificação dos conceitos presentes no assunto.
- Tradução desses conceitos nos termos de uma linguagem de indexação.

Em cada um desses estágios deve-se contar com a ajuda de instrumentos de indexação, como tesouros, códigos de classificação, cabeçalhos de assunto etc. (ABNT NBR 12676, 1992, p. 3).

Nunes questiona o uso da palavra traduzir, preferindo utilizar representar: “Ocorre que traduzir é um verbo complicado no âmbito documentário, exigindo-nos cautela no seu uso. Por quê? Porque em todas as etapas do processo de tratamento temático da informação o que fazemos é representação e não tradução” (NUNES, 2015, p. 2).

A primeira etapa do tratamento temático da informação é a análise de assunto, assim compreendida a leitura de um documento com a finalidade de identificar os conceitos representativos do seu conteúdo, representando-os por meio da linguagem de indexação adotada, observados os preceitos da política de indexação da biblioteca ou sistema de informação a que foi incorporado o documento (NUNES, 2015, p. 3).

Catálogo de assunto (em inglês: *subject cataloguing*), termo que remete à catalogação, mas está mais próximo dos princípios da indexação, é considerada por alguns autores como sinônimo de indexação. Porém, como afirmam Fujita, Rubi e Boccato (2009, p. 39), “a catalogação de assuntos está essencialmente ligada à construção de catálogos de bibliotecas, e a indexação, à construção de índices de bibliografias em serviços de informação bibliográficos que produzem bases de dados”.

Fujita, Rubi e Boccato (2009, p. 26), a partir de Fujita (1989) e Gomes e Marinho (1984), descrevem três princípios básicos da catalogação de assuntos: princípio do uso (descrições feitas da forma usada pelo usuário); da entrada específica (entrada pelo termo específico e não pela classe à qual está subordinado); e da estrutura sindética (que estabelece mecanismos para o relacionamento de cabeçalhos seja por equivalência, hierarquia ou associação). Para Fiuza (1985, p. 257), catalogação de assunto é a disciplina ou conjunto de disciplinas que tratam da representação, nos catálogos de bibliotecas, dos assuntos contidos no acervo.

Dois aspectos importantes para a compreensão e diferenciação dos processos de catalogação de assunto e indexação são relacionados por Fujita, Rubi e Boccato:

[...] ambos derivam de processos intelectuais para determinação de assuntos que melhor representem o documento para sua posterior recuperação, seja por meio de índices ou por meio de catálogos; e a tendência atual dos catálogos em atuarem como bases de dados, até mesmo com disponibilização de textos completos (FUJITA; RUBI; BOCCATO, 2009, p. 35).

Fujita, Rubi e Boccato (2009, p. 35) adotam em seus estudos o termo indexação para designar o procedimento realizado pelo bibliotecário no ambiente da biblioteca. Esse profissional realiza tanto a representação descritiva dos documentos quanto a descrição temática, caracterizando o processo de catalogação juntamente com o de indexação.

Fujita, Rubi e Boccato (2009, p. 36) afirmam que o bibliotecário precisa compreender que deve atuar com um indexador, realizando a análise de assunto para compreender o documento, identificando e selecionando os conceitos que melhor representem seu conteúdo durante o tratamento temático da informação com a finalidade de preencher o campo de assunto nos formatos catalográficos.

A partir da revisão teórica e histórica apresentada neste capítulo percebe-se o grande investimento intelectual ocorrido no último século a fim de desenvolver sistemas, ferramentas, padrões e práticas que levaram à melhor representação da informação – e, conseqüentemente, à melhor recuperação dessa informação. Também torna-se claro o quanto processos imprescindíveis em contextos “analógicos” precisaram se readequar para continuar a serem usados em cenários digitais. Em coleções volumosas como as das hemerotecas digitais, práticas como a indexação perdem espaço para o uso de tecnologias que consigam recuperar informação de forma rápida e satisfatória em textos completos. Ao mesmo tempo, processos como a catalogação, e sua tarefa de tornar um documento único, permanecem como imprescindíveis no refinamento de buscas realizadas em jornais ou revistas.

3 INFORMAÇÃO NO CONTEXTO DO JORNALISMO, A NOTÍCIA

A informação é a principal matéria-prima no Jornalismo, comumente denominada informação jornalística. Tal informação reflete um recorte espaço-temporal da realidade, sendo uma espécie de narração do tempo presente – com limitações, como não abordar a realidade em sua totalidade.

Extraída da realidade cotidiana, a informação jornalística apresenta-se em forma de nota, notícia, entrevista, reportagem etc. Desses gêneros, a notícia destaca-se como sinônimo do Jornalismo, um elemento que fundamenta propriamente a atividade, como comenta Quiroga ao resgatar temas abordados por Park em 1950.

Segundo Park, a prática profissional que se consolida na cena social norte-americana no início do século XX não poderia ser pensada senão a partir da centralidade da notícia, ou seja, seria ela, em última instância, o que caracteriza e define o ofício do jornalista. Em sua opinião, a notícia não seria apenas um gênero do Jornalismo, mas o elemento indissociável que definiria sua natureza e função, tornando-se a partir de então impossível falar de jornalismo sem necessariamente se remeter à ideia de notícia e vice-versa (QUIROGA, 2012, p. 370).

Lage (2001, p. 24) define: “notícia é a articulação simbólica que transporta a consciência do fato a quem não o presenciou”, complementando com: “Notícias são fragmentos de aparências” (LAGE, 2001, p. 35). Para o especialista, notícia também é o registro momentâneo de uma ocorrência, desde que esta seja atual, singular e extraordinária.

“O melhor da notícia é o retrato instantâneo, do qual se deve evitar ao máximo o relato descontextualizado, mas que seja o mais ágil possível em relação ao transcorrer do fato” (MOURA, 2012, p. 332). A autora destaca que no processo de padronização proposto pelo modelo norte-americano do Jornalismo, o texto da notícia foi, crescentemente, tornando-se um texto menos autoral e de aparência mais objetiva, ou seja, uma forma de relato mais distanciado, baseado ainda na imparcialidade.

“O Jornalismo quando narra os acontecimentos transforma-os em algo novo e, necessariamente, diferente do próprio acontecimento: a notícia (ou a reportagem, nota, o artigo de opinião, a entrevista etc.)”, como aponta Moura (2012, p. 329). Nessa visão, a notícia seria uma forma de relato. O processo de produção da notícia então, permanentemente sob exame, estaria situado entre dois pontos: da seleção à configuração de um fato em notícia. “Nesse intervalo, reside a prática jornalística, o ‘fazer’, a operação da construção da notícia”, explica Moura.

O acontecimento e seu relato nunca podem ser idênticos, dado que, ao relatar, o jornalista recorta uma parte da realidade, enfoca um ponto de vista, recolhe alguns testemunhos, retira este ou aquele trecho da fala do entrevistado, resume e relaciona este ou aqueles dados (MOURA, 2012, p.329).

Lage reúne diversas definições clássicas para a palavra notícia. Algumas, como descreve, ingênuas e genéricas. Nenhuma capaz de determinar de maneira única seu objeto.

- a) “Se um cachorro morde um homem, não é notícia; mas se um homem morde um cachorro, aí, então, a notícia é sensacional” (Amus Cummings);
- b) “É algo que não se sabia ontem” (Turner Catledge);
- c) “É um pedaço do social que volta ao social” (Bernard Voyenne);
- d) “É uma compilação de fatos e eventos de interesse ou importância para os leitores do jornal que a publica” (Neil MacNeil);
- e) “É tudo o que o público necessita saber; tudo aquilo que o público deseja falar; quanto mais comentário suscite, maior é seu valor; é a inteligência exata e oportuna dos acontecimentos, descobrimentos, opiniões e assuntos de todas as categorias que interessam aos leitores; são os fatos essenciais de tudo o que aconteceu, acontecimento ou ideia que tem interesse humano” (Colliers Weekly);
- f) “Informação atual, verdadeira, carregada de interesse humano e capaz de despertar a atenção e a curiosidade de grande número de pessoas” (Luís Amaral); (LAGE, 2001, p. 26).

Para Moraes (2002, p. 1), a informação jornalística preenche lacunas dos sistemas informacionais sobre diversos temas. “Há casos em que ela é a única fonte informacional”, avalia a autora, para em seguida complementar afirmando que a informação jornalística serve também para a preservação da memória nacional, local, institucional e, por vezes, pessoal.

A informação jornalística proporciona, também, uma consciência social maior que viabiliza a formação da ideia de identidade nacional através de extensão das interlocuções e principalmente dos temas e notícias pertinentes a um grupo social relatados por uma mesma língua (MORAES, 2002, p. 7).

Ao escrever uma notícia, o jornalista busca ser claro, preciso, direto, objetivo e conciso. A ordem direta e a simplicidade caracterizam o texto. Afinal, este precisa ser entendido por diferentes públicos. Há décadas servem de referência cinco questões principais: O quê? Quem? Quando? Onde? Por quê?

Garcia questiona: Como escreve um jornalista? Em seguida responde: Depressa. “A qualquer velocidade, exige-se fidelidade a três requisitos: exatidão (para não enganar o leitor), clareza (para que ele entenda o que lê) e concisão (para não desperdiçar nem o tempo dele nem o espaço do jornal)”, ressalta Garcia (1993, p. 15).

Darnton (2010, p. 42), historiador e bibliotecário que trabalhou como jornalista no início da carreira, argumenta que notícias sempre foram um artefato, e nunca corresponderam exatamente ao que realmente aconteceu. “Notícias não são o que aconteceu, mas uma história sobre o que aconteceu.” Pela experiência em uma redação, Darnton desconfia do jornal como fonte de informação. Para ele, “jornais devem ser lidos em busca de informações a respeito de como os acontecimentos eram interpretados pelas pessoas da época, em vez de representarem fontes confiáveis dos acontecimentos em si”.

Como observa Cavalcante (2002, p. 4), sempre haverá os que duvidam do valor dos jornais de época como fonte de pesquisa histórica alegando a interferência das ideologias no

ato de noticiar. Porém, assim como ideologias perpassam as páginas de qualquer jornal, sabe-se que ideologias não interferem apenas na veiculação de notícias jornalísticas, já que integram todo processo de produção e divulgação de ideias, em todos os tempos e lugares. Para a autora, estranho seria se jornais fossem totalmente isentos. “Tais veículos de informação acabam por espelhar múltiplas ideologias em confronto porque estão sempre a trazer conflitos existentes naquele meio social específico onde são produzidos” (CAVALCANTE, 2002, p. 4). Nesse contexto, cabe ao pesquisador identificar as ideologias presentes nas publicações selecionadas.

Não se pode negar que a imprensa escrita, formada por jornais e revistas, tem atributos que a qualificam como fonte potencial, sobretudo porque veicula um discurso e o materializa em formas textuais (carta dos leitores, editorial, artigos de opinião, textos noticiosos, classificados etc.); iconográficas (fotografias, charges, cartuns, gráficos, tabelas); e gráficas (distribuição hierárquica de textos e imagens na página).

O jornal como fonte histórica potencial é um lugar-espço social onde ocorreram e ocorrem relações diversas alimentadas pela ação social de quem, direta e indiretamente, está presente na notícia e na produção do jornal — apuração dos fatos, redação do texto e materialização gráfica (edição, ilustração e diagramação na página). Como tal, imprime o cotidiano de uma cidade, de um país, do globo (RIBEIRO; SILVA; SILVA, 2014, p. 228).

Assim, o jornal é o primeiro rascunho da história, o registro de pontos de vista sobre como os fatos transcorreram. “O jornal é como uma variada vitrine (de parte) da realidade reconstruída pelo fato jornalístico” (MOURA, 2012, p. 325). Além disso, o jornal é retrato do dia a dia – mesmo que um retrato parcial. Sobre este importante meio de comunicação se abordará mais detalhadamente no próximo capítulo.

Como afirma Moraes (2002, p. 7), o jornal é o espelho da sociedade em que está inserido. “Além de informar, o jornal emociona, diverte, agrupa socialmente e utiliza em suas narrativas elementos de polos dialéticos caracterizados pelas relações bom x mau, fraco x forte.” A autora pontua ainda que nesse cenário se evidencia que a ideia de conflito é base de interesse tanto para a produção quanto para a recepção de seu discurso.

O jornal é um bom referencial para se conhecer uma cidade. Nele estão expostos gostos, interesses, grau cultural, ideologia predominante, história cotidiana, tradição do grupo. O relato de acontecimentos, as seções especializadas, a coluna social, o obituário, os classificados, a publicidade, a forma de enunciação e ilustração das notícias produzidas pelo jornal espelham ele próprio e a sociedade que lhe atribui essas tarefas. No jornal está registrado o cotidiano que faz a história coletiva. As notícias publicadas nas páginas são histórias que criam a História. Mas nem todos os fatos ocorridos em uma sociedade são históricos. O que transforma um fato em acontecimento é ser conhecido, divulgado e registrado, principalmente escrito, lhe atribui credibilidade, é um passaporte para a história (MORAES, 2002, p. 10).

Desde que historiadores passaram a se interessar pelo cotidiano das pessoas comuns, não apenas pelos registros oficiais disseminados pelo Estado, os jornais ganharam espaço como fonte de informação. Essa “renovação historiográfica” deriva, segundo Calonga (2012, p. 80), das proposições advindas com a chamada *Nouvelle Histoire*, na década de 1970, que ampliou o entendimento de fonte documental na História.

Pesquisadores passaram a considerar a sociedade como um produto de diversos acontecimentos, signos e significados, que formam a cultura, o caráter e, conseqüentemente, a história. “O jornal virou importante fonte de informação ao tirar o pesquisador do isolamento científico e o inserir num espectro informacional maior, mais abrangente, sobre várias partes da realidade vistas por diversos relatores e apresentadas simultaneamente” (MORAES, 2002, p. 16). Por isso, o jornal torna-se fonte para a pesquisa social não apenas por mostrar a realidade, mas também por fazer parte de tal realidade.

Destaca-se na fala dos pesquisadores sempre a necessidade da leitura crítica do que está sendo obtido em jornais e revistas, análise criteriosa que o historiador faria em qualquer documento que fosse usar em suas pesquisas.

O próximo subcapítulo inicia-se com a evolução da imprensa até chegar ao jornal contemporâneo.

3.1 AS NOTÍCIAS IMPRESSAS, O JORNAL CONTEMPORÂNEO

Instrumentos e tecnologias ligadas à impressão já vinham sendo estudados e desenvolvidos há décadas em diferentes partes do mundo, mas foi por volta de 1450 que se encontrou a combinação perfeita para a reprodução em série. A “descoberta” é atribuída ao alemão Johannes Gutenberg, nascido e residente em Mainz (Mogúncia).

Como relembra Lage (2001, p. 9), a tecnologia gráfica resultou do comércio asiático. O papel era fabricado na Espanha, no século XII, sob ocupação árabe, e na Itália, no século XIV. Os tipos de metal surgiram na Coreia em 1390 e os de cerâmica datavam de dois séculos antes, criados por Pi Cheng em 1040, na China. O impulso seguinte foi dado com o início da colonização na América e a expansão do comércio com o Oriente.

Gutenberg foi quem conseguiu reunir todas essas tecnologias e fazer o processo de impressão se tornar conhecido na Europa. Ele utilizou tipos móveis forjados a chumbo e um instrumento inspirado nas prensas de uvas. A partir disso, os panfletos escritos à mão foram gradualmente sendo substituídos por folhas impressas, como as distribuídas em Veneza, no século XVI, que ganharam o nome de gazetas por custarem o valor de uma moeda, a *gazetta*.

Os primeiros periódicos com formato semelhante ao que se entende hoje por jornal circularam na primeira metade do século XVII.

Os exemplares mais antigos de jornal que se conhece foram publicados na Alemanha em 1609 e, embora não contenham indicações sobre a cidade ou o impressor, provavelmente saíram de uma oficina de Bremen. Outra publicação primitiva era editada em Estrasburgo (também em 1609) e uma terceira em Colônia (1610). Em 1620, havia jornais em Frankfurt, Berlim, Basileia, Hamburgo, Amsterdã e Antuérpia. Da Holanda saíram exemplares em inglês e francês nesse mesmo ano. Em 1621, circulava o primeiro jornal londrino, a *Current of General News*. Dez anos depois, a 30 de maio de 1631, era lançada em Paris *La Gazette*, de Theophraste Renaudot (LAGE, 2001, p. 11).

A exemplo da remota *Acta Diurna*, que publicava apenas as ações bem-sucedidas de Julio César, no século XVII era comum os jornais noticiarem somente glórias de seus apoiadores. Derrotas militares só se fossem as de outros países. A prática de noticiar fatos locais do cotidiano teve início da segunda metade em diante do século XVII, mas as informações eram frequentemente alvo de censores. Não se podia abordar nada que incitasse o povo a uma atitude de oposição ao governo. A primeira lei protegendo a liberdade de imprensa surgiu apenas em 1766 na Suécia.

Importante na evolução dos jornais foi a Revolução Industrial, iniciada em 1760, pois inseriu nesse setor a mecanização de processos de produção.

A impressora mecânica, inventada pelo alemão Koenig, foi utilizada pela primeira vez em 28 de novembro de 1814, na impressão do *Times*, de Londres; onze anos antes, entrava em operação a primeira máquina contínua para a fabricação de papel. Em 1867, Hippolyte Marinoni construía a prensa de quatro cilindros, a rotativa, cujo primeiro modelo era 25 vezes mais rápido do que qualquer outra máquina então existente. Mergenthaler inventou em Baltimore a linotipo que, a partir de 1880, aceleraria vitalmente a composição. A fotografia, inventada por Daguerre em 1839, abriu caminho para a fotogravura e a imprensa ilustrada. Data de 4 de março de 1880, a primeira reprodução de uma fotografia em jornal, no *Daily Graphic*, de Nova York (LAGE, 2001, p. 14).

Depois da invenção de Gutenberg, outra grande transformação na produção de jornais ocorreria por volta de 1844, com a criação do telégrafo. O sistema permitia a transmissão de informações em minutos, proporcionando a jornalistas notícias mais atuais e relevantes. A esta época, o jornal tornava-se o veículo mais importante de divulgação e recebimento de informações.

Não apenas em termos tecnológicos evoluíram os jornais, mas também a maneira de noticiar os fatos do dia a dia. Até a Revolução Industrial, as notícias eram relatos de acontecimento importantes no comércio, nos meios políticos e nas manufaturas. Com a conquista de um público leitor, como descreve Lage (2001, p. 24), as notícias tornaram-se artigos de consumo, sujeitos a acabamento padronizado, embalados conforme técnicas bem estudadas de marketing.

Para Sodré, a história da imprensa é a própria história do desenvolvimento da sociedade capitalista.

O controle dos meios de difusão de ideias e de informações – que se verifica ao longo do desenvolvimento da imprensa, como reflexo do desenvolvimento capitalista em que aquele está inserido – é uma luta em que aparecem organizações e pessoas das mais diversas situações social, cultural e política, correspondendo a diferenças de interesses e aspirações (SODRÉ, 1999, p. 1).

No cenário mundial, entre 1890 e 1920, formaram-se influentes impérios editoriais, como os construídos pelo norte-americano William Randolph Hearst (1863-1951), pelo húngaro radicado nos Estados Unidos Joseph Pulitzer (1847-1911) e pelo britânico Alfred Harmsworth, mais tarde conhecido como Lorde Northcliffe (1865-1922).

Com o surgimento de uma nova mídia, o rádio, nos anos 1920, os jornais tiveram de reavaliar seu papel. Novos formatos e conteúdos foram desenvolvidos para tornar os jornais mais atraentes, com coberturas mais amplas e maior profundidade.

Outras mídias, como a televisão e a “Web” impactariam diretamente os jornais nas décadas seguintes. Apesar das perdas de leitores e assinantes, os jornais continuam atingindo um bilhão de leitores diariamente, de acordo com dados da Associação Mundial de Jornais (WAN).

Machado de Assis escrevia em 1878 na revista *O Cruzeiro* que “o jornal, *literatura quotidiana*, no dito de um publicista contemporâneo, é reprodução diária do espírito do povo, o espelho comum de todos os fatos e de todos os talentos, onde se reflete, não a ideia de um homem, mas a ideia popular, esta fração da ideia humana”.

Nos anos 1930, Eric W. Allen definiu em *Journalism Quarterly*, a partir de uma adaptação a obra de Otto Groth (*Die Zeitung – Ein System des Zeitungskunde*) de 1928, que um jornal de verdade precisa ter sete características: deve ser periódico, reproduzível de maneira mecânica e estar disponível para quem possa pagá-lo, além de ser variado, geral, oportuno e organizado, como relatam Martin e Copeland (2003, p. 2).

O jornal, uma publicação periódica, como afirmam Guinchat e Menou (1994, p. 44), é um documento que aparece em volumes ou fascículos sucessivos, a intervalos mais ou menos regulares: são as coleções de obras, os relatórios periódicos, as revistas e os jornais. O conteúdo de cada edição é diferente. Guinchat e Menou (1994, p. 44) afirmam ainda que revistas e jornais são conhecidos como periódicos porque são produzidos em intervalos definidos e regulares. A periodicidade varia de 24 horas a um ano.

O jornal apresenta a realidade: a realidade vista pela lente do jornalismo e decodificada pelo código de tratamento da informação peculiar à área. Tem seu tempo diferenciado, sua visão específica, suas cores e ritmo especiais. Nisto, reside a sua arte, recortar a realidade e emoldurá-la com palavras, títulos, fotos, legendas e gráficos. Também aí reside o seu grande risco: o recorte pode ser parcial demais; o ângulo de visão pode ser o mais deformado; a fonte entrevistada pode ser a menos comprometida com a verdade dos fatos; a fala recortada pode não ser a mais significativa. Toda perícia e seriedade na apuração e produção da notícia fazem-se necessárias nesse entroncamento (MOURA, 2012, p. 330).

Quanto à apresentação física do jornal, considerando título, fotos, diagramação de colunas e outras características estilísticas, esta costuma ser constante.

A forma do jornal é a primeira pista para o entendimento de seu lugar na cultura contemporânea, a compreensão de sua linguagem e a investigação de sua história. A disposição das manchetes, o desenho das letras, sua uniformidade ou variedade, a existência ou não de claros e o equilíbrio estético entre eles, o tamanho e a natureza das ilustrações poderão nos informar se o jornal se destina a público mais ou menos amplo, de menor ou maior escolaridade. Cada detalhe nos remete a uma categorização: o modo como se distribuem os elementos gráficos (a *paginação* ou *projeto gráfico*) relaciona-se com escolas e correntes de arte, de modo que alguns jornais se enquadram no design industrial despojado e outros lembram a organicidade flamejante do *art nouveau* (LAGE, 2001, p. 9).

Moura (2012, p. 326) afirma que a própria hierarquização das chamadas de primeira página do jornal sinaliza ao leitor a quais fatos está sendo atribuída maior ou menor relevância por meio de determinado relato jornalístico.

De acordo com Pena (2013, p. 66), a primeira tentativa de classificação no jornalismo foi feita pelo editor Samuel Buckeley no começo do século XVIII quando decidiu separar o conteúdo do jornal *Daily Courant* em *news* (notícias) e *comments* (comentários). Ao longo dos séculos seguintes, estudos na área seguiram o critério da separação entre forma e conteúdo, gerando divisão por tema, assim como pela relação do texto com a realidade, ou seja informação x opinião.

No Brasil, como revisto por Pena (2013, p. 67), o assunto foi pesquisado por Luiz Beltrão e José Marques de Melo, cujas propostas baseavam nos seguintes critérios: 1) Finalidade do texto ou disposição psicológica do autor; 2) Estilo; 3) Modos de escrita ou morfologia ou natureza estrutural; 4) Natureza do tema e topicalidade; e 5) Articulações interculturais (cultura).

Marques de Melo inventariou algumas das principais classificações (Quadro 2), mostrando o quanto são marcadas pela diversidade e sujeitas a críticas de quem as considera inapropriadas.

Quadro 2 – Classificações dos textos que compõem um jornal

Classificação francesa
(autor: Joseph Foliet)
<p> Editorial Artigos de fundo Crônica geral (resenhas dos acontecimentos) Despachos (reportagens e entrevistas) <i>Fait divers</i> Crônica especializada (crítica) Folhetim Fotos e legendas Caricaturas <i>Comics</i> (quadrinhos) </p>
Classificação norte-americana
(autor: Fraser Bond)
<p> Noticiário notícia reportagem entrevista história de interesse humano Página editorial editorial caricatura coluna crítica </p>
Classificação alemã
(autor: Emil Dovifat)
<p> Informativos notícia (<i>fact-story</i>) <i>report</i> (<i>act-story</i>) entrevista (<i>quote-story</i>) De opinião editorial artigos curtos glosa (crônica) Amenos folhetim (resenha cultural) crítica recreio e espelho cultural (contos, versos etc.) </p>

Fonte: Pena (2013, p. 67-68).

A proposta de Marques de Melo a partir do trabalho de Beltrão, como recuperado por Pena (Quadro 3), desenha-se da seguinte forma:

Quadro 3 – Proposta de classificação de Marques Beltrão

Jornalismo informativo
nota
notícia
reportagem
entrevista
Jornalismo opinativo
editorial
comentário
artigo
resenha
coluna
crônica
caricatura
carta

Fonte: Pena (2013, p. 68).

Este é um tema controverso, mas de modo geral, as editorias pouco variam de um jornal para outro. Até mesmo os nomes utilizados são semelhantes: opinião, Brasil/país, economia, mundo/internacional, segundo caderno/variedades ou algo equivalente, caderno de esportes ou equivalente, revista de domingo, caderno de tecnologia, cadernos especiais, como de gastronomia, moda e veículos.

“O *menu* não se altera muito, mas se comporta com estrutura diferente se falamos de um jornal de circulação nacional ou de um jornal regional/local” (MOURA, 2012, p. 326). A variação está relacionada não apenas à questão regional/nacional do veículo, mas ao perfil dos leitores, a tiragem e ao próprio plano editorial do jornal.

Os conteúdos apresentados neste capítulo reforçam o valor e a importância da informação jornalística como um retrato da realidade – e conseqüentemente possível fonte tanto para atender a necessidades prosaicas de informação quanto para a utilização em investigações científicas. Como descrito anteriormente, a informação reflete um recorte espaço-temporal, sendo uma espécie de narração do tempo presente, mesmo que com limitações. A notícia, melhor expressão do jornalismo, baseia-se em cinco questões principais, as mesmas que servem de suporte e até mesmo inspiração aos leitores que pretendem não apenas informar-se sobre fatos cotidianos, mas investigar comportamentos, características e absorver todas as nuances do conteúdo apresentado no texto noticioso.

4 DA NECESSIDADE À RECUPERAÇÃO DA INFORMAÇÃO

A falta do necessário para compreender uma situação, tomar uma decisão, realizar uma pesquisa ou sanar uma dúvida leva um indivíduo a buscar informações. No âmbito da Ciência da Informação, necessidades como essas desencadeiam o processo de busca e recuperação da informação.

As necessidades de informação traduzem um estado de conhecimento no qual alguém se encontra quando se confronta com a exigência de uma informação que lhe falta e lhe é necessária para prosseguir um trabalho. Ela nasce de um impulso de ordem cognitiva, conduzido pela existência de um dado contexto (problema a resolver, objetivo a atingir) e pela constatação de um estado de conhecimento insuficiente ou inadequado (MIRANDA, 2006, p. 102).

Cole (2011, p. 1216) destaca que a necessidade de informação denota tradicionalmente o estágio inicial de alguém em busca de uma resposta para sua questão. Na abordagem dominante, como descreve o autor, adotada especialmente pela Ciência da Computação, a formulação da consulta (*query*⁶) a um Sistema de Recuperação da Informação requer determinado tipo de atitude do usuário. Este deve, de alguma forma, coletar suas reflexões sobre a informação necessária, identificar e rotular esses pensamentos em termos e, em seguida, prever quais deles irão obter os resultados que atenderão à necessidade informacional.

Cole, porém, prefere não seguir cegamente a teoria dominante. Para o autor (2011, p. 1216), diferente da necessidade de comida, água ou proteção, o que é requerido para satisfazer uma necessidade de informação é frequentemente desconhecido pelo usuário. Por isso o contexto ou a situação que leva o usuário a ter esta necessidade torna-se tão importante. Cole toma como base os estudos desenvolvidos por Taylor, especialmente o artigo de 1968, no qual o autor apresenta um modelo de necessidade de informação em quatro níveis, sendo o nível mais profundo aquele que é intangível e não especificado.

Nesse estágio, como escreve Taylor (1968, p. 268), há um vago sentimento de insatisfação, mas difícil de ser expresso em termos linguísticos. A necessidade, ainda não identificada propriamente, irá mudar de forma, qualidade, concretude e de critérios à medida que o indivíduo for tomando melhor conhecimento da dúvida que lhe aflige.

No segundo nível o autor trata sobre uma descrição mental consciente de uma área de indecisão mal definida, que resultará provavelmente em uma questão ambígua e desconexa. A fim de aguçar o foco, o indivíduo poderá conversar com alguém sobre o problema. Nesse caso, podem ocorrer duas situações: a) a pessoa com quem ele compartilhou

⁶ *Query* é o construto-chave para a recuperação da informação (JANSEN; RIEH, 2010, p. 1528).

seus pensamentos entenderá a ambiguidade ou b) as ambiguidades irão gradualmente desaparecer no decurso do diálogo.

No terceiro nível, o indivíduo consegue formular uma questão racional e qualificada. Ele descreve a área em que tem uma dúvida em termos concretos e pode ou não pensar dentro de um contexto ou restrições de um sistema do qual ele quer a informação.

No quarto nível, a questão é formulada por antecipação ao que os mecanismos podem entregar. O consulente precisa pensar baseado na organização do mecanismo em particular que pretende usar e nos materiais que estarão disponíveis, como livros, jornais, relatórios, artigos, desenhos ou tabelas.

Taylor resume os quatro níveis da seguinte forma:

- Q1 – a atual, mas não expressada necessidade de informação, a necessidade visceral.
- Q2 – a consciente, dentro do cérebro descrição da necessidade, a necessidade consciente.
- Q3 – a questão formal da necessidade, a necessidade formalizada.
- Q4 – a questão como apresentada para o sistema de informação, a necessidade comprometida (TAYLOR, 1968, p. 255, tradução nossa).

Taylor descreve uma situação em que o indivíduo com uma necessidade de informação chega a uma unidade de informação. Há que se considerar Taylor dentro do contexto em que estava inserido no final dos anos 1960. Naquela época, o comportamento padrão era entregar ao bibliotecário algumas palavras-chave (Q4). Entrava em cena a habilidade do bibliotecário para retirar desse indivíduo um pouco mais de informações a fim de ajudá-lo a encontrar o que realmente precisava. Esperava-se que este profissional conseguisse retroceder ao nível Q2, para então percorrer o caminho, junto com o indivíduo, até chegar ao Q4.

Este é um processo direto e estruturado, embora existam, é claro, muitos diferentes estilos e muitos níveis de competência e conhecimento tanto da parte do indivíduo que busca a informação quanto do bibliotecário. Existem traços óbvios que ajudarão os bibliotecários: empatia, capacidade de se colocar no lugar do outro, conhecimento do assunto, conhecimento sobre a coleção e a clientela (TAYLOR, 1968, p. 255).

Na pesquisa desenvolvida por Taylor para escrever o artigo referenciado foram entrevistados bibliotecários e profissionais da informação. A partir do material coletado, o autor elaborou uma lista com os cinco itens mencionados pelos entrevistados que mais os ajudaram a chegar à real necessidade de informação dos indivíduos que os procuraram, aspectos que ajudam a refinar a estratégia de busca.

A lista é composta por:

1. Delimitação e estruturação do assunto a ser buscado. Assim como um aprofundamento sobre a área. Tudo isso irá ajudar a refinar, expandir, aprofundar

e qualificar o processo de busca.

2. Objetivos e motivação. Por que o indivíduo quer determinada informação? Quais são seus objetivos? Qual a sua motivação? As respostas podem ajudar a qualificar ainda mais o assunto ou até mesmo alterar a busca. Além de ajudar a estimar qual o tipo de resposta será o mais adequado? Taylor (1968, p. 257) escreve que a menos que saibamos qual é o porquê, nunca teremos certeza do que a pessoa realmente quer ou o que fará com a informação. Não poderemos ajudar o indivíduo a menos que saibamos sobre suas necessidades tanto quanto ele.
3. Características pessoais do indivíduo com uma necessidade de informação. O que ele estuda? A que organização está ligado? Já consultou o acervo em questão? Qual a sua experiência em buscas do tipo que pretende realizar? Qual a relação do indivíduo com aquilo que está buscando? Respostas a estas perguntas ajudam no processo de negociação entre o indivíduo com o bibliotecário. A partir das respostas pode-se determinar a urgência da busca e estimar como os resultados serão aceitos.
4. Relação da descrição da questão com a coleção. O processo de negociação entre o indivíduo com a necessidade de informação e o bibliotecário oferece uma descrição completa sobre a questão de busca, assim como pistas que ajudam a elaborar a estratégia de busca. O bibliotecário poderá então atuar como um tradutor, interpretando e reestruturando a questão para que se encaixe no que poderá ser recuperado na unidade de informação.
5. Respostas previstas e aceitáveis. Quando o indivíduo consulta uma unidade de informação, ele costuma ter em mente uma ideia do que vai encontrar, como será, qual o volume de dados recuperados. No processo de negociação, o bibliotecário procura ter uma ideia dessas expectativas do indivíduo – e alterá-la caso esteja muito fora das capacidades tanto da unidade de informação quanto da estratégia de busca.

Apesar de o contexto ser um tanto diverso do cenário atual, no qual o indivíduo vê-se sozinho diante do sistema de recuperação da informação, a reflexão elaborada por Taylor continua bastante pertinente e útil àqueles que vão empreender uma busca. Por essa razão, a lista desenvolvida por Taylor servirá como base para os estudos empíricos realizados neste trabalho. Mesmo que os cinco passos descritos anteriormente não sejam fixos e precisem de adequação a cada caso, eles mostram-se práticos na organização e clarificação do complexo processo de definição da necessidade de busca.

4.1 CONTEXTO NA CIÊNCIA DA INFORMAÇÃO

Na palavra contexto estão contidas as condições inter-relacionadas em que algo existe ou ocorre⁷. Contexto pode ser entendido como a moldura que circunda uma determinada situação. Moldura esta que influenciará na maneira como a situação em questão será interpretada⁸.

Como escreveram Capurro e Hjørland em seu tratado sobre o conceito de informação, definições não são verdadeiras ou falsas, mas sim, mais ou menos produtivas. “Os diferentes significados dos termos que usamos são ferramentas mais ou menos eficientes para ajudar-nos a alcançar o que pretendemos” (CAPURRO; HJØRLAND, 2007, p. 151). Assim, no presente trabalho, a palavra contexto será entendida como o ambiente no qual algo ocorre, e que precisa ser considerado para melhor compreensão do que está sendo relatado.

Na Ciência da Informação, seu objeto maior de estudos, a informação está sempre inserida em um contexto específico – seja a informação contida em documentos impressos, em bibliotecas ou objetos, seja aquela da comunicação informal, de repositórios digitais ou presente na “Web”. Saracevic (1999, p. 1054) afirma que a informação em Ciência da Informação envolve não apenas uma mensagem que será processada cognitivamente, mas um contexto, ou seja, uma situação, uma tarefa, um problema.

Capurro e Hjørland afirmam que “a informação não é uma propriedade de fatos, mas é dependente do contexto e das limitações” (2007, p. 170). Na ampla revisão empreendida pelos autores sobre o conceito de informação, ganham destaque temas relacionados à assimilação da informação, ao conhecimento prévio e às diferentes influências e interpretações da informação observadas em diversas áreas do conhecimento, das Ciências Exatas às Ciências Humanas, passando pela Ciência da Informação.

Sob determinado ponto de vista, Capurro e Hjørland apontam a perspectiva do receptor, que envolve crenças, desejos e contexto próprios, e é bastante considerada no processo informacional. Em outra interpretação, os conteúdos de informação não são dependentes do conhecimento do receptor, mas dos tipos de situação. O significado é determinado nos contextos social e cultural. “Dois receptores diferentes podem extrair, devido a limitações e fatos diversos, diferentes conteúdos de informação a partir de um mesmo sinal” (CAPURRO; HJØRLAND, 2007, p. 171).

⁷ Definição do dicionário *Merriam-Webster*: “the interrelated conditions in which something exists or occurs: ENVIRONMENT, SETTING <the historical *context* of the war>”. Disponível em: <<http://www.merriam-webster.com/dictionary/context?show=0&t=1342807112>>. Acesso em: 15 jul. 2012.

⁸ Definição apresentada no portal *Comunicação & Comportamento*. Disponível em: <<http://www.portalcmc.com.br/glossarioc.htm>>. Acesso em: 15 jul. 2012.

No desenvolvimento de suas considerações, Capurro e Hjørland citam Ludwig Wittgenstein, mais especificamente os jogos de linguagem, os quais, de certa forma, referem-se a uma espécie de contexto composto por palavras e interpretações. Ainda na revisão de Capurro e Hjørland, contexto está presente na perspectiva do gerenciamento de conhecimento no âmbito da chamada Sociedade da Informação. Peças isoladas de dados significantes constituem conhecimento quando integradas a um contexto.

Por fim, Capurro e Hjørland citam Hobart e Schiffman, os quais afirmam não ser a informação um fenômeno que surge com a tecnologia moderna, mas um produto de complexas interações entre tecnologia e cultura. “Hobart e Schiffman distinguem entre as eras clássica, moderna e contemporânea da informação, sendo que o significado da informação é específico em cada uma” (CAPURRO; HJØRLAND, 2007, p. 174), apontando a influência do contexto.

Além de Saracevic, Capurro e Hjørland, outros autores trataram sobre contexto – Wersig e Nevelling, em 1975, e Belkin e Robertson, em 1976. Deve-se lembrar que esses quatro autores destacam a responsabilidade social dos cientistas da informação, os quais devem atuar para que o conhecimento seja transmitido para aqueles que precisam dele, como afirmam Wersig e Nevelling (1975, p. 134), uma vez que a Ciência da Informação é uma disciplina orientada para a solução de problemas.

Wersig e Nevelling apresentam definições e usos da informação a partir dos contextos de quem a está utilizando, de acordo com as necessidades informacionais suscitadas em cada grupo. Já Belkin e Robertson, inspirados por Wersig e Nevelling, procuram delimitar seções específicas (contextos) dentro do espectro da informação com as quais a Ciência da Informação deve se preocupar, sendo essas seções em nível infracognitivo, cognitivo individual, cognitivo-social e metacognitivo.

Brookes (1980) lança um olhar sobre os espaços em que as informações são tratadas. O autor toma como ponto de partida os três mundos de Popper: o mundo 1, composto por matéria, energia e radiação; o mundo 2, por nossos pensamentos, sentimentos e imagens; e o mundo 3, por conhecimento objetivo, livros, publicações, registros, artefatos criados pelo homem.

Brookes observa haver uma deficiência nas definições do filósofo. Popper descreve os objetos e entidades que compõem cada um de seus mundos, mas, como considera Brookes, diferentes mundos não têm apenas diferentes “mobiários”, talvez possam ter diferentes espaços. “Se não considerarmos a possibilidade que diferentes mundos têm diferentes tipos de espaço, tenderemos a deduzir que todos os espaços são idênticos àqueles que exploramos em

detalhe – aqueles do espaço físico” (BROOKES, 1980, p. 129). Pensar melhor sobre os espaços presentes no mundo 3, segundo Brookes, é um dos desafios dos cientistas da informação.

4.2 BUSCA DA INFORMAÇÃO E RECUPERAÇÃO DA INFORMAÇÃO

Tanto a área de busca da informação (*information seeking*) quanto a de recuperação da informação (*information retrieval*) focam na interação entre pessoas e o conteúdo de sistemas de informação.

Tradicionalmente, a busca de informação tende a ser um tema de bibliotecários e cientistas da informação, como visto no item anterior, enquanto que a recuperação da informação está mais ligada a profissionais da Ciência da Computação e áreas afins, segundo Jansen e Rieh (2010, p. 1517). Os estudos desenvolvidos na área de recuperação da informação têm como foco descobrir como melhorar o desempenho tecnológico do sistema.

A essência da busca de informação está na formulação de uma questão a ser submetida a um sistema de informação, a qual fará esse sistema trabalhar, como explica Cole (2011, p. 1217). A questão apresenta-se em forma de uma consulta (*query*).

A *query* é um conjunto de um ou mais símbolos que combinado com a sintaxe (a disposição dos símbolos) é usado como um comando em um SRI para localizar possíveis conteúdos relevantes indexados no SRI em questão. Símbolo em uma consulta pode ser uma palavra, uma frase, um grupo de palavras, um documento, uma imagem ou muitas outras possibilidades, como definem Jansen e Rieh (2010, p. 1528).

Antes de chegar à *query* em si, o indivíduo costuma passar por algumas etapas – ou deveria – que lhe auxiliam a aumentar a recuperação em uma busca, como visto nos estudos desenvolvidos por Taylor. Definida a consulta, parte-se para a busca e – no transcorrer ideal do processo – para a recuperação da informação.

Busca, de acordo com o *Dicionário de Biblioteconomia e Arquivologia*, está relacionada à recuperação da informação, ou seja, a:

1. Restituição dos dados constantes do sistema, para obtenção de informações específicas ou genéricas. A restituição, ou recuperação, abrange o processo total de identificação, busca, encontro e extração da informação armazenada. Nesta operação não se incluem nem a criação, nem a utilização posterior das informações ou dos dados; restituição da informação.
2. Recuperação de dados, informação ou documentos de uma coleção ou acervo a partir de um pedido formulado.
3. Ação, métodos e procedimentos que têm por objetivo extrair de um conjunto de documentos as informações desejadas.

4. Área que estuda as técnicas de tratamento, armazenamento e recuperação da informação em um arquivo ou acervo de documentos, com o objetivo de atender às necessidades de informação de um usuário (CUNHA; CAVALCANTI, 2008, p. 307).

Recuperação significa extração de informação de uma coleção de conteúdos, como definem Jansen e Rieh (2010, p. 1517). O termo recuperação da informação foi cunhado por Calvin Mooers em 1948. A ideia da recuperação da informação já era sugerida por Vannevar Bush alguns anos antes. Ao apresentar o *memex* no artigo *As we may think*, Bush descreve hipoteticamente como seria acessar grandes quantidades de conhecimento armazenado no equipamento, além de dar uma ideia de como funcionaria a lógica de pensamento do usuário, ou seja, não de forma linear, mas por meio de conexões entre temas.

Não foi somente Mooers que tratou do assunto em 1948. Na *Royal Society Scientific Information Conference* desse mesmo ano um número expressivo de artigos mencionava o comportamento informacional (*information behavior*), ou seja, o comportamento dos sujeitos ao buscar por informação. O termo ainda não era usado nos artigos, que tratavam essencialmente sobre uso de documentos e bibliotecas, mas a origem da pesquisa sobre o tema estava claramente nessa época, como escreve Wilson (1999, p. 250).

O final dos anos 1950 e o começo dos 1960, de acordo com Lesk (1995, p. 4), foram marcados por experimentos em SRI, que funcionavam como interface entre o usuário e os documentos que compunham uma coleção. As discussões geradas pelos testes na década de 1960, nas quais era questionada a indexação de todas as palavras em um documento e não somente das palavras-chave, levaram ao surgimento de medidas de avaliação, como revocação e precisão, desenvolvidas por Cyril Cleverdon, do então *Cranfield College of Aeronautics*, como aponta Lesk (1995, p. 5).

Segundo Cleverdon, revocação (*recall*) refere-se à fração de documentos recuperados em relação ao total, enquanto que precisão diz respeito à fração de documentos recuperados que são relevantes.

Na Ciência da Informação, as definições de revocação e precisão foram disseminadas por Lancaster. Para o autor (2004, p. 4), revocação é a capacidade de recuperar documentos úteis em uma base de dados. Uma busca com “revocação alta” será aquela em que a maioria dos itens pertinentes, se não todos, serão recuperados. O coeficiente de revocação – medida da extensão com que se alcança a precisão desse processo – é o número de itens pertinentes recuperados dividido pelo número total de itens.

Precisão, para Lancaster (2004, p. 4), é a capacidade de evitar documentos inúteis. O coeficiente de precisão – medida da extensão com que se alcança a precisão do processo de recuperação relevante – é o número de itens relevantes recuperados dividido pelo número total de itens recuperados. Morville (2005, p. 49) resume de forma simplificada o conceito de relevância, precisão e revocação: “Resultados relevantes são aqueles interessantes e úteis para os usuários. [...] Precisão mede o quão bem um sistema recupera ‘apenas’ documentos relevantes. Revocação mede o quão bem um sistema recupera ‘todos’ os documentos relevantes”.

Retomando o histórico da busca e da recuperação da informação nos anos 1960: etapa importante ocorreu no dia 29 de outubro de 1969 quando cientistas norte-americanos estabeleceram uma rede de transmissão entre quatro computadores instalados em diferentes pontos dos Estados Unidos – na Universidade da Califórnia, Los Angeles (UCLA), na Universidade da Califórnia em Santa Bárbara, na Universidade de Utah e no Instituto de Pesquisa de Stanford. Era dado o primeiro passo rumo ao desenvolvimento da rede mundial de computadores.

A partir da década de 1970, computadores realizavam cálculos complexos, gravando dados para serem usados posteriormente. Por conta disso, diferentes testes dos sistemas de recuperação *off-line* puderam ser realizados. No sistema manual de busca de uma biblioteca, composto por fichas catalográficas, era praticamente impossível usar operações booleanas como *and*, *or* e *not*. Lancaster (2004, p. 50) observa que as entradas lineares desses sistemas não permitiam as buscas de informações pela combinação de palavras. Já na busca em uma base de dados no computador havia como recuperar materiais que mencionassem dois assuntos em particular.

Nos anos 2000, Morville apresenta o termo *findability* (“encontrabilidade”) para tratar sobre a busca e recuperação da informação. Para o autor, *findability* requer “definição, distinção e diferença”, ou seja, tudo que está inserido em uma coleção ou base de dados precisa ser representado de maneira que se torne distinguível de outros itens em um acervo – o que remete ao princípio básico da catalogação. Em ambientes físicos, isto pode ser aplicado por meio de tamanhos, formas, cor e a localização dos objetos separadamente. Já no reino digital, é necessário confiar nas palavras – seja em forma de rótulos, de *links*, mas, especialmente, no formato de palavras-chave.

A humilde palavra-chave tornou-se surpreendentemente importante nos últimos anos. Como um ingrediente vital no processo de busca *on-line*, palavras-chave são inseridas no *Google*, *Yahoo!*, *MSN*, *eBay* e *Amazon*. A busca de notícias, produtos, pessoas, móveis usados e música. E as palavras são a chave para o nosso sucesso (MORVILLE, 2005, p. 4).

Certamente nem tudo é tão simples quanto parece. Afinal, como o próprio Morville observa, as palavras são “criaturinhas bagunçadas”, que variam de significado de acordo com o contexto, podendo representar o paraíso para um usuário de um sistema de busca e o inferno para outro. Sinônimos, antônimos, homônimos são verdadeiros desafios aos responsáveis pela representação da informação – nesse cenário observa-se a importância dos tesauros e vocabulários controlados.

O termo palavra-chave, segundo Guinchat e Menou (1994, p. 145), é empregado para designar as palavras extraídas diretamente do conteúdo do documento, da linguagem natural. Elas, que foram desenvolvidas a partir da automação das unidades de informação, costumam ser usadas da mesma maneira que se apresentam.

As palavras-chave são escolhidas empiricamente nos documentos em razão de sua aptidão em expressar o conteúdo das informações. Elas permitem uma descrição muito aprofundada e são combináveis entre si infinitamente. São reagrupadas em listas alfabéticas. A dificuldade de sua manipulação é compensada pela utilização do computador. Entretanto, é difícil assegurar a coerência deste tipo de linguagem, o que leva a fazer certas escolhas, isto é, selecionar algumas palavras-chave que se tornam descritores (GUINCHAT; MENO, 1994, p. 145).

Como escrito anteriormente, “buscar algo que você conhece (um endereço, um número de telefone) exige um comportamento muito diferente daquele de quando você está em busca de algo desconhecido”, afirmam Morville e Rosenfeld (2006, p. 31). Distinguir entre essas duas necessidades e esses dois comportamentos apontando quais são as principais prioridades dos usuários é uma tarefa extremamente valiosa. Os resultados dessa avaliação determinam com maior precisão onde investir esforços, recursos, tempo e dinheiro ao desenhar a arquitetura de informação do site.

Ao longo dos anos 1960 e 1970, foram desenvolvidos os primeiros sistemas de busca em texto integral. Os experimentos, como relata Morville (2005, p. 50), usavam o Sistema de Armazenamento e Recuperação da Informação da IBM. Tanto a precisão quanto a revocação alcançavam números impressionantes, entre 75% e 80%. Logo se passou a alardear o fim dos metadados. Para que gastar tempo e dinheiro contratando humanos para indexar termos quando os sistemas de recuperação de texto integral funcionavam tão bem? A automação parecia ser o futuro.

Porém, vale lembrar que “metadados indicados por humanos tratam sobre a atinência (*aboutness*) do documento, o que melhora a precisão.” (MORVILLE, 2005, p. 53) Segundo Morville, os mecanismos do *Google* amam metadados, considerando-os excelentes indicadores para saber mais sobre o que trata determinado documento existente na “Web”. Infelizmente, o esforço para a etiquetagem manual de conteúdos, devido ao alto custo de

tempo e dinheiro, torna-a proibitiva para a maioria dos acervos volumosos. “Por isso, este recurso nem sempre consegue ser usado exatamente onde mais precisava”, observa Morville (2005, p. 53).

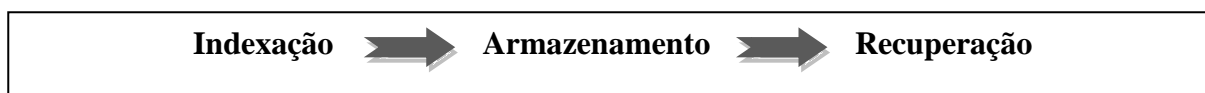
Os pesquisadores, que estavam prestes a dispensar os metadados, só se esqueceram de levar em conta o tamanho das coleções quando analisaram os números obtidos nas pesquisas. Quando o volume de materiais era pequeno, os resultados eram satisfatórios. O mesmo não ocorria em grandes acervos. Os sistemas recuperavam apenas um quinto dos documentos relevantes – e somente quando eram utilizados por usuários experientes. Provou-se que a máquina tinha muitas utilidades, mas o trabalho do profissional da informação continuava (e continua) imprescindível.

4.3 SISTEMAS DE RECUPERAÇÃO DA INFORMAÇÃO

Qual o propósito de um sistema de recuperação da informação, o que ele deve fazer? Robertson (1981, p. 10) responde: “Levar o usuário até aqueles documentos que melhor satisfaçam a sua necessidade de informação”.

De acordo com Rowley (2002, p. 162), os sistemas de busca e recuperação da informação estão relacionados a três processos de trabalho ao lidar com as informações (Figura 5).

Figura 5 – Processos de trabalho dos SRI com a informação



Fonte: Rowley (2002, p. 162).

Vieira (2014, p. 164) reforça que para que haja uma recuperação rápida e sem ruídos deve existir um armazenamento adequado das informações no sistema. Afinal, como já escrito, só se consegue recuperar o que foi bem representado. Para isso, é necessário que se tenha feito um bom trabalho de indexação das informações, para que o consultante não perca tempo na recuperação de informações genéricas em sua busca, a não ser que seja esse seu objetivo durante esse processo.

Para Souza, os SRI organizam e viabilizam o acesso aos itens de informação. Para tanto, desempenham atividades determinadas, como:

- a) Representação das informações contidas nos documentos, normalmente por uso de processo de indexação e descrição dos documentos;
- b) Armazenamento e gestão física e/ou lógica dos documentos e de suas representações;
- c) Recuperação das informações representadas e dos próprios documentos armazenados, de forma a satisfazer as necessidades de informação dos usuários (SOUZA, 2006, p. 163).

Nesse último item, a interface e os recursos oferecidos são essenciais, pois será por meio deles que os usuários irão inserir suas necessidades e, após receber os resultados, filtrar e analisar o que mais lhes interessa.

Miranda ressalta que a função de um SRI é a possibilidade de busca e recuperação, facilitando ou permitindo o acesso à informação. “Seu propósito é atender a uma necessidade de informação. Portanto, um SRI não poderá ser avaliado sem considerar seu propósito e sua função” (MIRANDA, 2005, p. 40).

Rowley apresenta o processo de busca dividido em três etapas, que poderão ser repetidas pelo usuário várias vezes até que fique satisfeito com o resultado:

- Aceitação de uma consulta como insumo (como uma representação da necessidade de informação), formulada pelo usuário.
- Execução de uma comparação da consulta com cada um dos registros (representação dos documentos) existentes na base de dados.
- Produção, como resultado a ser submetido ao usuário, de um conjunto de registros recuperados e que foram identificados com base nessa comparação (ROWLEY, 2002, p. 164).

A partir da revisão, que contempla período de 20 anos, entre os anos 1970 e 1990, Rowley elaborou um quadro (Quadro 4) em que apresenta as diferentes gerações dos SRI.

Quadro 4 – Gerações dos SRI

Geração	Tipo	Característica
Primeira geração	Metadados	Interfaces baseadas em comandos, usuários especialistas e intermediários; número limitado de sistemas <i>on-line</i> nas instituições, e disponíveis externamente por meio de serviços de busca <i>on-line</i> .
Segunda geração	Dados com texto integral	Interfaces baseadas em menus e comandos; recursos de recuperação adicionais, como hipertextos e buscas em texto completo; interfaces baseadas em DOS; previsto o acesso pelo usuário final, mas nem sempre possível ou alcançado; sistemas <i>on-line</i> , com os primeiros sistemas apresentados em CD-ROM.
Terceira geração	Multimídia	Interfaces gráficas; foco no acesso realizado pelo usuário final; orientada para o mercado e com ênfase em pacotes de produtos; armazenamento e distribuição em CD-ROM ou em redes de alta capacidade; multimídia; intermediário com a função de instrutor; maior uso no lar e em ambientes de acesso público.

Fonte: Rowley (2002, p.167).

Mais de 20 anos se passaram desde o levantamento realizado por Rowley. O cenário sofreu transformações significativas a partir da metade dos anos 1990 com a chegada da internet comercial à vida das pessoas. Surgiram então os primeiros mecanismos de busca na rede. Os serviços pioneiros oferecidos eram compostos por catálogos e diretórios, construídos manualmente utilizando-se de categorias temáticas. *Yahoo!*, *Lycos*, *HotBot* e o brasileiro *Cadê* foram alguns dos pioneiros. No *Yahoo!*, por exemplo, as informações sobre tubarões estavam classificadas da seguinte forma: *Ciência > animais, insetos e bichos de estimação > Vida aquática > Peixes > Espécies > Tubarões*.

As informações nos diretórios de assunto, como descreveria Branski (2004, p. 73) à época, eram coletados de duas formas:

- a) Busca realizada pelos editores dos sites, que visitam inúmeros sites e incluem os de interesse no banco de dados, acompanhados de uma breve descrição de seus conteúdos;
- b) Solicitação de inclusão enviada pelo autor interessado em ter seu site catalogado. O autor enviava uma breve descrição do conteúdo, e os editores visitavam o site, aceitando ou não sua inclusão.

Em pouco tempo, o modelo tornou-se inviável, pois o volume de informações crescia em progressão geométrica. Popularizaram-se os programas de busca, não mais oferecendo dados inseridos no sistema de forma manual, mas com o auxílio de *spiders* ou robôs.

Os robôs percorrem a rede coletando informações. Iniciam seu caminho a partir de um conjunto de páginas selecionadas por seus administradores que são escolhidas por serem populares, de alta qualidade ou por conterem grande quantidade de *links*. Os robôs visitam os sites, lendo seu conteúdo, armazenando cada uma das palavras encontradas e seguindo seus *links* para outras páginas. Seguem os *links* para encontrar outros *links* e ir, sucessivamente, adicionando os endereços ao banco de dados. Diferentemente dos catálogos, os sites são incluídos no banco de dados sem nenhuma classificação ou descrição de seu conteúdo (BRANSKI, 2004, p. 74).

O modelo de organização pelo qual passavam os acervos de bibliotecas e as bases de dados não foi levado para os grandes sites de busca que se tornavam muito populares na virada do século XX, como *AltaVista* e *Google*. Cada mecanismo fornecia diferentes resultados, pois também o acervo variava de um site para outro, dependia da capacidade dos robôs na coleta de informações. Desde os primórdios da rede, a busca e recuperação da informação já se apresentavam como questões complexas.

Como afirma Eco (2012, p. 48), o problema maior da “Web” – e estendendo-se isso a qualquer outro grande acervo – é que encontrar o que se quer depende da capacidade de quem consulta. O escritor não se refere somente aos mecanismos de busca, mas ao verdadeiro desafio contemporâneo relacionado à competência em informação: não adianta apenas saber

usar os recursos de busca e recuperação da informação, importante também é ter discernimento para separar o que é relevante entre tudo o que se recupera.

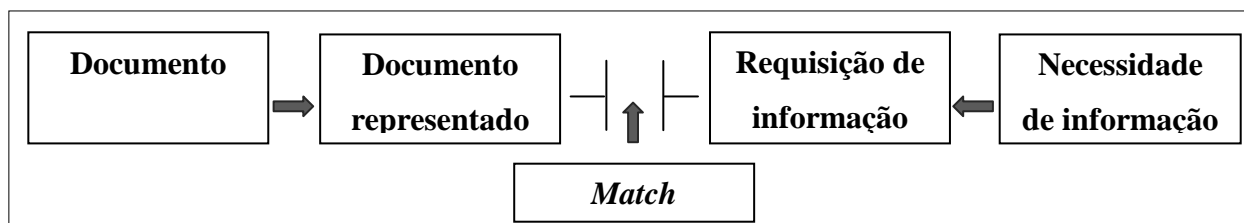
Yang – que considera buscar informação na “Web” algo diverso em caráter e imprevisível por natureza – refere-se à diversidade de usuários, com seus diferentes *backgrounds* e motivados pelos mais variados tipos de necessidades informacionais. Não são mais apenas profissionais habilitados que lidam com os sistemas. “A variedade de experiências, conhecimento, motivação e propósitos significa que os usuários podem expressar diferentes tipos de necessidades de informação por meio de grande heterogeneidade de formas, com diferentes critérios, para satisfazer suas necessidades” (YANG, 2005, p. 33).

Sobre os “novos” usuários de SRI, Manning, Raghavan e Schütze (2009, p. 432) observam que na tradicional recuperação da informação usuários eram tipicamente profissionais com treinamento sobre como criar consultas para uma coleção cujo estilo e estrutura eles conheciam bem; já os usuários da “Web” tendem a não saber – ou se importar – com a heterogeneidade do conteúdo e a sintaxe das linguagens de criação de consulta. “Ademais, nem é papel de uma plataforma como a “Web” exigir esse tipo de conhecimento de bilhões de pessoas” (MANNING; RAGAVHAN; SCHÜTZE, 2009, p. 432).

Com o advento da “Web” e dos mecanismos de busca, “buscar informação se tornou ubíquo”, como afirma Saracevic (2010, p. 162). O renomado cientista da informação afirma que “todo mundo é um buscador hoje em dia”. Essa nova realidade se instalou, entre outras razões, devido aos avanços em recuperação da informação, que englobam, por exemplo, os grandes mecanismos de busca contemporâneos.

Nas origens dos SRI, o processo de busca e recuperação da informação costumava ser representado pelo seguinte modelo (Figura 6).

Figura 6 – Processo de busca e recuperação da informação



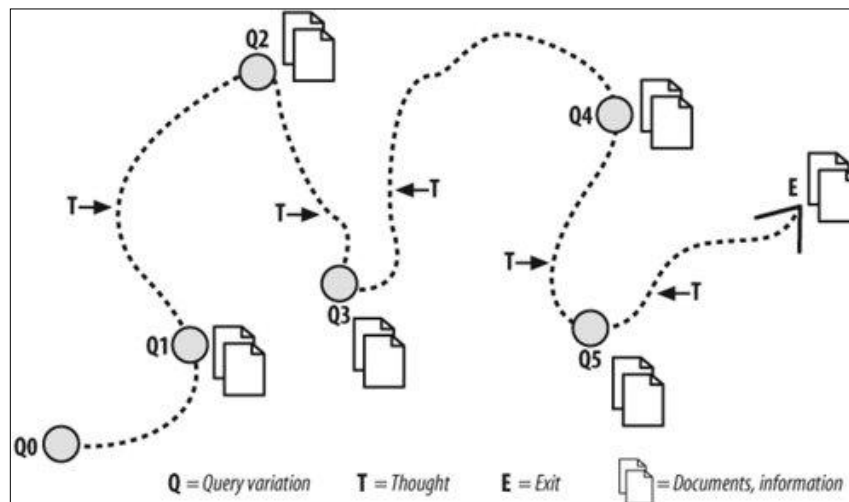
Fonte: Bates (1989, p. 408).

Bates (1989, p. 408) afirma que tal modelo foi muito produtivo e ajudou no entendimento do processo de recuperação da informação. Porém, como apontado por Thomas Kuhn, modelos considerados centrais em um determinado campo – e em determinado

momento – começam a mostrar inadequações e insuficiências à medida que são aplicados testes para entender o processo que está sendo estudado. As limitações tornam-se assim evidentes e aos poucos buscam-se novas soluções, estabelecendo-se um novo paradigma.

Como estudou Bates, na vida real os usuários começam suas buscas com um tema amplo. De acordo com os resultados obtidos, vão alterando e melhorando a busca até chegar ao resultado que melhor atenda a sua necessidade. A partir de diversas experiências de busca seguindo esta ideia, Bates apresentou um novo modelo, o qual chamou de *Berry-picking model* (Figura 7), pois sua representação se assemelhava a uma pessoa colhendo amoras em diversos arbustos. À medida que vai realizando as buscas e obtendo determinados resultados, o sujeito altera suas palavras-chave e a direção do caminho que estava percorrendo.

Figura 7 – Modelo *Berry-picking*



Fonte: Bates (1989, p. 410).

Os SRI costumam oferecer mais comumente duas opções de busca: a busca simples e a busca avançada. A primeira se caracteriza por um campo em branco no qual o usuário escreverá as palavras-chave que considera melhor relacionadas à sua necessidade informacional. O exemplo mais conhecido deste tipo de busca é a página do *Google*, composta por um campo de busca em destaque.

A busca avançada pode ser apresentada de diversas formas, já com opções para refinamento da busca e com uma utilização mais ampla de linguagem documental. Os operadores booleanos costumam aparecer em forma de expressões como “com todas as palavras” ou “com alguma das palavras”. Há buscas avançadas que utilizam filtros – como no caso da Hemeroteca Digital da Biblioteca Nacional brasileira, na qual são utilizados os filtros local, título e período, além da possibilidade de se escrever palavras-chave.

Apesar dos esforços de pesquisadores em apontar o comportamento “ideal” dos usuários, estudos realizados por diversos cientistas, como apontam Aula e Nordhausen (2006, p. 1678), mostram que os usuários típicos da “Web” escolhem apenas, em média, dois termos de busca, realizam sessões de busca curtas, checam apenas uma página de resultados e normalmente cometem erros ao usar a busca avançada. Há, portanto, um longo caminho para quem estuda o comportamento dos usuários de SRI.

4.4 LÓGICA E ESTRATÉGIA DE BUSCAS

Rowley define lógica de buscas como o “meio de especificar combinações de termos que devam ser obtidas para se chegar a uma recuperação bem-sucedida” (2002, p. 171).

Uma das lógicas mais utilizadas é a booleana, que serve para ligar os termos que descrevem conceitos presentes no enunciado da busca. Os operadores booleanos surgiram no século XIX, criados pelo matemático George Boole (1815-1864).

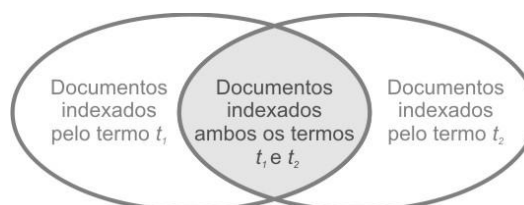
A chamada álgebra booleana é um sistema binário no qual existem somente dois valores possíveis para qualquer símbolo algébrico: 1 ou 0, verdadeiro ou falso. Conforme a revisão de Ferneda (2003, p. 22), no modelo booleano, um documento é representado por um conjunto de termos de indexação que podem ser definidos de forma intelectual (manual) por profissionais especializados ou automaticamente por meio da utilização de algum tipo de algoritmo computacional.

As buscas são formuladas por meio de expressão booleana composta por termos ligados através dos operadores lógicos *and* (Figura 8), *or* (Figura 9) e *not* (Figura 10) (e, ou e não, respectivamente) e apresentam como resultado os documentos cuja representação satisfaz às restrições lógicas da expressão de busca.

Ferneda resume o funcionamento dos três operadores em uma busca:

A expressão conjuntiva de enunciado t_1 *and* t_2 recuperará documentos indexados por ambos os termos (t_1 e t_2). Esta operação equivale à interseção do conjunto dos documentos indexados pelo t_1 com o conjunto dos documentos indexados pelo termo t_2 , representado pela área de interseção na Figura 8.

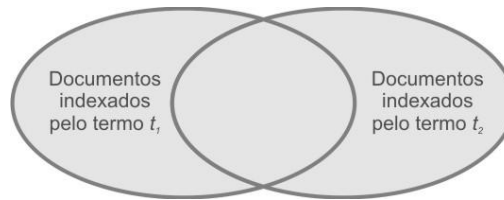
Figura 8 – Representação do resultado de expressão booleana conjuntiva (*and*)



A expressão disjuntiva t_1 *or* t_2 recuperará o conjunto dos documentos indexados pelo termo t_1 **ou** pelo termo t_2 . Essa operação equivale à união entre os conjuntos dos

documentos indexados pelo termo t_1 e o conjunto dos documentos indexados pelo termo t_2 , não havendo interseção de conteúdos, como ilustrado na Figura 9.

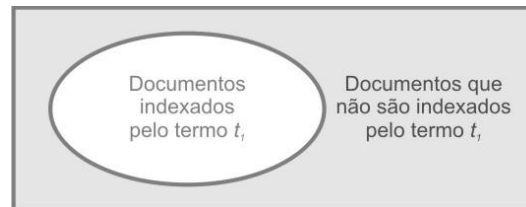
Figura 9 – Representação do resultado de uma busca booleana disjuntiva (*or*)



A expressão que utiliza apenas um termo t_1 terá como resultado o conjunto de documentos indexados por t_1 . A expressão *not* recuperará os documentos que não são indexados pelo termo t_1 , representados pela área cinza da Figura 10.

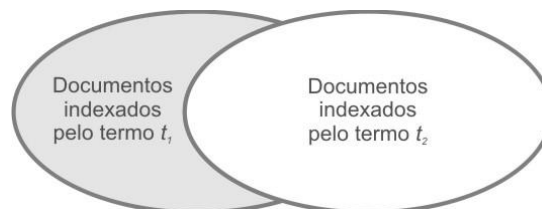
10

Figura 10 – Resultado de busca negativa (*not*)



As expressões t_1 *not* t_2 terão o mesmo resultado: o conjunto dos documentos indexados por t_1 e que não são indexados por t_2 (Figura 11). Neste caso, o operador *not* pode ser visto como um operador da diferença entre conjuntos. Assim a área cinza da figura 11 representa o conjunto dos documentos indexados pelo termo t_1 menos (subtraído de) o conjunto dos documentos indexados por t_2 .

Figura 11 – Resultado de busca booleana com o operador *not*



(FERNEDA, 2003, p. 22-23).

Como já considerado, quando o usuário sabe o que quer e onde buscar, a consulta a um acervo é simples. Por exemplo, procurar por um telefone no site de uma empresa ou buscar a população de determinado país em uma base de dados estatísticos – locais onde as informações costumam ser apresentadas de forma mais estruturada.

A busca da informação em um ambiente ainda pouco conhecido pelo usuário pode diminuir a certeza de que encontrará o que procura, assim como dificultar a definição de como irá procurar determinada informação. Se o ambiente não se apresentar de forma organizada, rapidamente o usuário procurará uma alternativa ou desistirá da busca.

Vale lembrar também que:

É um mito pensar que o usuário sempre planeja a estratégia de busca ou que faz buscas muito elaboradas ou avançadas. Uma busca geralmente corrobora o princípio do menor esforço; o usuário deseja que o sistema recupere informação com qualidade e que ele tenha facilidade na utilização. Esse comportamento do usuário se aplica tanto à biblioteca convencional quanto à digital (CUNHA, 2008, p. 8).

A dificuldade, como dito anteriormente, intensifica-se quando a informação de que se necessita está armazenada em um site pouco estruturado ou se o usuário não sabe exatamente que tipo de palavra-chave escolher, qual recurso de busca utilizar ou mesmo qual vai ser a estratégia a ser adotada.

Estratégia de busca, segundo Lopes (2002, p. 60), é uma técnica ou um conjunto de regras para tornar possível o encontro entre uma pergunta formulada e a informação armazenada em uma base de dados.

Bates (1979), pioneira no estudo de estratégias e táticas de buscas, especialmente pensadas para serem realizadas em bases de dados, antecede a conceituação revisando os significados dos termos estratégia e tática, ressaltando que ambos são mais conhecidos por seus usos militares.

De forma geral, estratégia lida com um plano mais amplo, enquanto tática refere-se a objetivos em curto prazo, podendo ser útil para alguns casos, mas não para todos. Bates dividiu as táticas de busca de informação em quatro categorias: táticas de monitoramento (que englobam todo o processo de busca); de arquivos estruturados (uso de bases de dados estruturadas para encontrar informação); de formulação da busca (como os termos reunidos para uma busca, considerando também o uso de operadores booleanos); e de escolha de termos de busca.

A partir disso, Bates (1987, 1998 apud LOPES, 2002, p. 62) conceitua estratégia de busca como o estudo da teoria, princípios e prática de planejar e executar tais táticas. Assim, o planejamento e a execução de uma estratégia de busca é a arte de escolher onde, quando e com que investigar cuidadosamente a fonte de informação para alcançar os objetivos específicos que desencadearam a necessidade de busca.

Smith (2011) retoma, pouco mais de 30 anos depois, o texto de Bates testando e adequando as táticas para a busca de informação na “Web”, tendo como objetivo ajudar tanto os usuários quanto os profissionais de informação que lidam com educação e competência em informação. O resultado é uma lista com 34 táticas (Quadro 5), sendo 18 originalmente pensadas por Bates, estas marcadas com sublinhado.

Quadro 5 – Táticas de busca da informação

Tipo de tática	Tática	Definição
Monitorar	<u>Checar</u> (<i>check</i>)	Rever a consulta original e compará-la com a consulta corrente para ver se é a mesma.
	<u>Ponderar</u> (<i>weigh</i>)	Realizar avaliação de custo-benefício, em um ou mais momentos da busca, de ações em curso ou das que virão.
	<u>Padronizar</u> (<i>pattern</i>)	Estar atento a um padrão de busca, examinando-o e redesenhando se não for maximamente eficiente ou se estiver desatualizado.
	<u>Corrigir</u> (<i>correct</i>)	Observar e corrigir erros de digitação e fatuais.
	<u>Gravar</u> (<i>record</i>)	Acompanhar caminho percorrido e caminhos desejados que não foram percorridos ou completados.
Estrutura do arquivo (índices, catálogos, resumos)	<u>Bibliografia</u> (<i>bibble</i>)	Verificar se já não há uma bibliografia pronta, antes de começar a elaborar uma. De forma geral: verificar se trabalho de busca já não foi feito e usado por outros.
	Prover (<i>provider</i>)	Visitar sites que têm mais probabilidade de prover informação; usar um site como uma faceta na busca.
	Endereço eletrônico (URL = localizador padrão de recursos)	Entrar URL para encontrar um site; usar URL ou nome do domínio na busca; modificar URL de uma página relevante para obter informações adicionais.
	Sistema centro-periferia (<i>hub & spoke</i>)	Seguir <i>links</i> de uma página de referências em um caminho centro-periferia, talvez usando diferentes janelas/abas ou o botão “voltar”.
	Encontrar (<i>find</i>)	Usar o recurso de busca do navegador (<i>control F</i>) para buscar texto em uma página.
	<i>Links</i> apontando para outros sites (<i>backlink</i>)	Buscar páginas que levam a uma página “ <i>pearl</i> ”; útil também para identificar o novo site para uma página que trocou a URL.
	Agregar valor (<i>value add</i>)	Buscar por subscrições ou bancos

		de dados de valor agregado que contenham material não indexado em mecanismos de busca. Isso pode incluir subtáticas: buscar informações em sites de bibliotecas etc.
	Mídias sociais (<i>socialmediate</i>)	Procurar informações em mídias sociais.
	Viajar no tempo (<i>timetravel</i>)	Usar arquivos em “ <i>cache</i> ” ou arquivos de internet para encontrar sites que não existem mais.
Formulação (formulation)	Teleportar (<i>teleport</i>)	Usar palavras-chave e um mecanismo de busca para ir diretamente ao objetivo.
	<u>Exaustividade</u> (<i>exhaust</i>)	Incluir a maioria ou todas as facetas da consulta na formulação inicial da busca; adicionar uma ou mais facetas da consulta a uma formulação já preparada.
	<u>Reduzir</u> (<i>reduce</i>)	Minimizar número de facetas da consulta na formulação inicial; subtrair uma ou mais facetas da consulta de uma formulação já pronta.
	<u>Paralelo</u> (<i>parallel</i>)	Fazer a formulação da busca ampliada incluindo sinônimos ou termos conceitualmente paralelos.
	<u>Localizar</u> (<i>pinpoint</i>)	Formular busca precisa por meio de redução ou minimização do número de termos paralelos, retendo os termos mais perfeitamente descritivos.
	<u>Bloquear</u> (<i>block</i>)	Para rejeitar, na formulação da busca, itens que contenham ou indexem determinados termos, mesmo se isto significar perder alguns documentos que tiverem itens relevantes.
	Sobreposição mínima (<i>minoverlap</i>)	Buscar nas facetas que tenham um mínimo de sobreposição, a fim de aproximar a busca o máximo possível.
Termo (term)	<u>Subordinar</u> (<i>sub</i>)	Mover para baixo hierarquicamente para um termo mais específico (subordinado).
	<u>Superordenar</u> (<i>super</i>)	Mover para cima hierarquicamente

		para um termo mais abrangente (superordenado).
	<u>Relativo</u> (<i>relate</i>)	Mover para o lado hierarquicamente para um termo coordenado.
	<u>Contrário</u> (<i>opposite</i>)	Buscar o termo logicamente oposto daquele que descreve a informação desejada.
	<u>Rastrear</u> (<i>trace</i>)	Examinar a informação já encontrada a fim de localizar termos adicionais a serem usados nas buscas subsequentes.
	<u>Reespaçar</u> (<i>respace</i>)	Tentar espaçar ou pontuar de maneira variante.
	<u>Ressoletrar</u> (<i>respell</i>)	Buscar com palavras soletradas de formas diferentes.
	Frase (<i>phrase</i>)	Usar frase na busca para maximizar ranking de termos composto por muitas palavras.
	Antecipação de termos (<i>anticiterm</i>)	Antecipar os tipos de termos que irão aparecer nas páginas relevantes.
Avaliação (<i>evaluate</i>)	Contexto (<i>context</i>)	Olhar outras páginas do site, a fim de estabelecer em que tipo de site estão inseridas as páginas recuperadas, observar a autoria do site etc.
	Verificar (<i>crosscheck</i>)	Verificar as informações em outras fontes.
	<i>Cachet</i> (<i>cachet</i>)	Usar ferramentas de busca, como diretórios, para pré-avaliar informações, ou procurar evidências de certificação do site.
	Audição (<i>audition</i>)	Usar a aparência da página – parte gráfica, design, qualidade do texto etc. – como um indicador de credibilidade.

Fonte: Smith (2011, p. 11).

O passo seguinte à inserção de uma *query* – ou seja, da requisição de informação, normalmente em linguagem natural, utilizando-se de palavras-chave – no sistema de busca costuma ser ou *searching* (pesquisar) ou *browsing* (navegar) ou *asking* (pedir/perguntar ajuda a alguém), ou, ainda, a combinação dessas alternativas.

O processo encerra-se com a listagem dos resultados da busca, ou seja, a recuperação da informação, a ser analisada pelo indivíduo com a necessidade informacional, processo que será desempenhado melhor ou pior de acordo com a sua competência em informação.

4.5 COMPETÊNCIA EM INFORMAÇÃO

Um mundo de informações descortina-se à frente dos usuários da “Web”, rico em possibilidades. Cabe a cada um balancear privacidade, liberdade, conveniência e segurança. Em meio a tantas possibilidades, cada um será colocado em teste.

“Como tomaremos nossas decisões? Como saberemos o suficiente para fazer as perguntas certas? A quem perguntaremos? Em quem confiaremos? Como faremos para encontrar o melhor produto, para a pessoa certa, o dado que fará a diferença?”, pergunta Morville (2005, p. 3), bibliotecário especialista em “encontrabilidade”.

Cunha (2008, p. 9) também lança um questionamento sobre o que uma unidade de informação, seja ela convencional ou digital, pode fazer para ajudar e aconselhar aqueles que têm dificuldades de avaliar o conteúdo e a utilidade das informações hospedadas na rede. Para Miranda (2006, p. 103), “as pessoas usam a informação para resolver problemas ou desenvolver uma tarefa. O ambiente social na qual a informação é encontrada determina seu valor e sua importância”.

No cenário composto por bilhões de páginas “Web”, utilizado por milhões de pessoas, identificar o que é relevante e útil entre os resultados obtidos em um processo de busca e recuperação da informação está relacionado à competência em informação, ou seja, às habilidades que uma pessoa competente em informação deve ter:

1. reconhecer a necessidade de informação;
2. reconhecer que informações acuradas e completas são a base para a tomada de decisões inteligentes;
3. formular questões baseadas na necessidade de informação;
4. identificar fontes de informação potenciais;
5. desenvolver estratégias de busca adequadas;
6. acessar fontes de informação inclusive as eletrônicas;
7. avaliar informações;
8. organizar informações para aplicações práticas;
9. integrar novas informações ao corpo de conhecimentos existentes;
10. usar informações para pensar criticamente e para solucionar problema (CAMPELO, 2009, p. 36).

De acordo com Amorim (2013, p. 27), a competência em informação compreende todo o processo de manejar a informação, ou seja, as habilidades de ler, usar, avaliar, selecionar e organizar informações. Tais capacidades permitem a comparação do processo de classificar, essencial na vida diária de qualquer indivíduo.

Como mostram as revisões de literatura realizadas pelas pesquisadoras Hatschbach (2002), Dudziak (2003) e Campello (2003), a expressão competência em informação foi registrada pela primeira vez em 1974 no relatório *The information service environment relationships and priorities* do bibliotecário Paul Zurkowski.

Zurkowski (1974, p. 6) destaca que indivíduos treinados na aplicação de fontes de informação em seu trabalho podem ser considerados como aqueles que aprenderam técnicas e habilidades para lidar com um grande número de ferramentas informacionais, bem como com fontes primárias, para encontrar informação que contribuam na solução de seus problemas.

Segundo a ALA (1989)⁹, “competência em informação é o conjunto de habilidades que o indivíduo deve ter para reconhecer quando uma informação é necessária e para localizar, avaliar e usar de forma efetiva informação requerida”.

Considerando a globalização e a cultura digital, Kong et al., em 2005, como referenciado por Campello (2009, p. 69), afirma que a pessoa competente na onipresente sociedade da informação deveria saber como determinar a natureza e a extensão da informação de que necessita, como encontrar informação efetiva e eficientemente e como usar informação para compreender questões econômicas, legais e sociais relacionadas com o uso da informação. Competência em informação, para Kong et al., é a habilidade de dominar os processos de ser informado e constitui capacidade essencial, necessária aos cidadãos para se adaptarem à cultura digital, à globalização e à emergente sociedade baseada no conhecimento.

A partir das considerações referenciadas anteriormente, Dudziak (2003, p. 29) apresenta os objetivos da competência em informação na formação de indivíduos que:

- Saibam determinar a natureza e a extensão de sua necessidade de informação como suporte a um processo inteligente de decisão;
- Conheçam o mundo da informação e sejam capazes de identificar e manusear fontes potenciais de informação de forma efetiva e eficaz;
- Avaliem criticamente a informação segundo critérios de relevância, objetividade, pertinência, lógica e ética, incorporando as informações selecionadas ao seu próprio sistema de valores e conhecimentos;
- Usem e comuniquem a informação, com um propósito específico, individualmente ou como membro de um grupo, gerando novas informações e criando novas necessidades informacionais;
- Considerem as implicações de suas ações e dos conhecimentos gerados, observando aspectos éticos, políticos, sociais e econômicos extrapolando para a formação da inteligência;
- Sejam aprendizes independentes; e
- Aprendam ao longo da vida (DUDZIAK, 2003, p. 29).

⁹ Definição de letramento informacional, sinônimo de competência em informação, da *American Library Association* (ALA). Disponível em: <<http://goo.gl/Q2ARsH>> Acesso em: 14 jun. 2015.

O usuário contemporâneo, diante dos inúmeros recursos tecnológicos disponíveis, precisa ter, segundo a ALA (2011)¹⁰, a habilidade para usar as tecnologias de informação e comunicação para encontrar, avaliar, criar e comunicar a informação, o que requer não apenas capacidade técnica, mas também cognitiva. No cenário do *big data*, composto por vastas bases de dados virtuais difíceis de serem processadas e analisadas, surgem novos desafios para os profissionais da informação, que precisam desenvolver novas habilidades para a competência em informação digital.

Ainda de acordo com a definição da ALA (2011), uma pessoa competente em informação digital:

- a) Tem uma variedade de habilidades – técnicas e cognitivas – requeridas para encontrar, entender, avaliar, criar e comunicar informação digital em uma ampla variedade de formatos;
- b) É capaz de usar diversas tecnologias de forma apropriada e de forma efetiva para recuperar informação, interpretar os resultados e julgar a qualidade da informação;
- c) Entende a relação entre tecnologia, aprendizado ao longo da vida, privacidade pessoal e administração da informação;
- d) Usa essas habilidades e a tecnologia apropriada para comunicar e colaborar com pares, colegas, família e o público em geral; e
- e) Usa essas habilidades para participar ativamente na sociedade e contribuir para uma vibrante, informada e engajada comunidade.

Ao buscar informações em uma unidade de informação, o usuário deve desenvolver a capacidade de refinar as consultas, assim como a de usar mecanismos de busca para descobrir outras opiniões ou comentários sobre o assunto recuperado.

Outra competência essencial é o chamado *Knowledge assembly*, em uma tradução livre “ajuntamento de conhecimento”, ou seja, a capacidade de extrair informações de diferentes fontes e formatos, seja um texto, um vídeo ou um tópico em uma lista de discussão, conseguindo avaliar o valor e as armadilhas de cada um. Só assim o indivíduo estará apto a tirar melhor proveito do que recuperar no processo de busca da informação.

Neste capítulo foram abordadas questões relacionadas à necessidade de informação. Desde o momento em que esta necessidade começa a ser delimitada pelo usuário até a

¹⁰ Definição de information literacy, sinônimo de competência em informação digital, da ALA. Disponível em: <<http://connect.ala.org/node/181197>> Acesso em: 25 jan. 2016.

inserção em um sistema de informação como os utilizados nas hemerotecas digitais. Nem sempre a necessidade apresenta-se de maneira pronta ou totalmente clara, daí a necessidade de prestar atenção a detalhes importantes como o contexto em que está inserida a fonte de informação ou mesmo o tema que causa incertezas nos usuário.

Escolhidos os recursos em um mecanismo de busca, estabelecidas as estratégias, resta ao leitor, no caso de jornais e revistas, ter discernimento e preparo suficiente para avaliar os resultados recuperados. A leitura crítica é obrigatória, especialmente se o conteúdo será usado em investigações científicas. Por meio desta revisão, buscou-se recuperar processos, cenários e habilidades que podem interferir diretamente na busca da informação que melhor atenderá a necessidade do usuário.

5 BIBLIOTECAS

Arqueólogos escavavam um palácio da antiga cidade de Ebla, na Síria, em 1975, quando descobriram um arquivo contendo fragmentos e partes maiores de 15 mil tabuinhas originais que ainda estavam em suas prateleiras originais quando foram incendiadas em 2300-2250 a.C. Como descreve Campbell (2015, p. 38), esta talvez não seja a biblioteca mais antiga, mas serve para mostrar como era a aparência física das bibliotecas da Mesopotâmia – com estantes por todos os lados e bancos de alvenaria encostados nas paredes.

Historiadores costumam considerar a biblioteca localizada em Nínive, antiga Assíria, às margens do Rio Tigre, como a mais antiga da humanidade, descoberta quase por acidente pelo arqueólogo amador e aventureiro Henry Layard em 1849. Na Biblioteca de Nínive, também conhecida como Biblioteca de Assurbanipal, foram encontrados mais de 25 mil tabletas de argila, nos quais há registros sobre religião, mágica, geografia, ciência, poesia, assim como sobre a história e o governo da época.

Estima-se que o rei Assurbanipal viveu entre 668 a.C. e 627 a.C. “Assurbanipal desempenhou a tarefa (de coletar materiais) com muita energia, buscando e adquirindo coleções, mas não se sabe se ele fez uso da força ou do dinheiro” (CAMPBELL, 2015, p. 40). Tem-se esta como a primeira tentativa de coletar o conhecimento de forma sistemática. Grande parte da coleção que restou de Nínive encontra-se hoje no Museu Britânico, em Londres.

A Biblioteca de Nínive pode ser a mais antiga, mas a mais famosa e referenciada certamente é a Biblioteca de Alexandria. Como escreve Manguel (1997, p. 215), Alexandre, o Grande fundou Alexandria em 331 a.C., mas a biblioteca, na qual se pretendia reunir todo o conhecimento do mundo, foi criada pelo sucessor de Alexandre, Ptolomeu I, influenciado pelo erudito grego Demétrio de Falero. Apesar da enorme fama da biblioteca, por falta de registros confiáveis nunca foi possível precisar o tamanho e a composição da coleção. Estima-se que tenha sido construída em 323 a.C., ou mesmo, em 246 a.C.

Contemporânea e concorrente da Biblioteca de Alexandria, a Biblioteca de Pérgamo foi fundada por Eumedes II – neto de Filetero, um dos generais de Alexandre, o Grande – onde hoje se encontra a cidade de Bergama, na Turquia. Ambas foram criadas no período helenístico, marcado pela difusão da cultura grega, compreendido entre 323 a.C. e 197 a.C. Assim como no caso da Biblioteca de Alexandria, não restaram registros que forneçam informações precisas sobre a coleção. Sabe-se apenas que parte substancial do acervo apresentava-se em rolos de pergaminho, tipo de suporte desenvolvido na região. Se o fogo destruiu a coleção alexandrina, em Pérgamo a umidade foi a grande inimiga da biblioteca.

Na Idade Média, que engloba a queda do Império Romano até o Renascimento (entre 600 e 1500 d. C.), as maiores bibliotecas não estavam na Europa cristã, mas no mundo árabe e no sudeste da Ásia. De acordo com Campbell (2015, p. 61), o papel permitiu aos estudiosos islâmicos – que haviam se espalhado do norte da África e pelo Oriente Médio, chegando até a Espanha – produzir livros em maiores quantidades e distribuí-los amplamente.

A tecnologia da produção de papel, no entanto, não foi inventada no Oriente Médio, mas na China, onde o uso de papel era combinado à impressão para produzir livros muito diferentes dos europeus, séculos antes de essa tecnologia estar disponível na Europa (CAMPBELL, 2015, p. 61).

Passaram-se séculos e os papéis das bibliotecas evoluíram. Atualmente, existem diferentes tipos de acervos, cada qual voltado para um público distinto. Bibliotecas podem ser classificadas de acordo com seus objetivos. Enquanto bibliotecas públicas visam à democratização da cultura, bibliotecas universitárias são responsáveis pelo apoio documental ao ensino e à pesquisa. Já as Bibliotecas Nacionais têm como meta a preservação e a difusão do patrimônio bibliográfico de cada nação, como explica Fonseca (2007, p. 8).

No subcapítulo seguinte, outras definições e características atribuídas a uma Biblioteca Nacional.

5.1 BIBLIOTECAS NACIONAIS

Monte-Mór (1972, p. 15) afirma que uma Biblioteca Nacional é “sinônimo da memória documental da cultura de um país, no seu sentido mais alto, museu de toda a sua produção bibliográfica, nos mais diversos campos culturais, através de sua história”.

A Biblioteca Nacional de um país é o seu repositório bibliográfico oficial e único. De valor incalculável, deve ser preservado de forma perene. É patrimônio de indiscutível interesse histórico-cultural, próprio e único de cada país – daí a relevância de seu acervo e, conseqüentemente, a importância dos procedimentos de organização e de representação da informação para descrição, uso e preservação (SOUZA, 2014, p. 11).

Utilizando como exemplo a Biblioteca Nacional do Brasil, Portella (2010, p. 5) afirma que uma biblioteca insere-se no conceito de nacional, em contraposição ao de pública, pelas seguintes características: ser beneficiária do instituto do depósito legal; possuir mecanismo estruturado para compra de material bibliográfico no exterior a fim de reunir uma coleção de obras estrangeiras nas quais se incluam livros relativos ao Brasil ou de interesse para o país; elaborar e divulgar a bibliografia brasileira corrente por meio de catálogos *online*; ser o centro nacional de permuta bibliográfica, em âmbito nacional e internacional.

Na explicação de Portella, destaca-se uma característica essencial na constituição do acervo das Bibliotecas Nacionais: o depósito legal,

definido como a exigência, por força das leis nº. 10.994, de 14/12/2004 e nº 12.192, de 14/01/2010, de remessa à Biblioteca Nacional de um exemplar de todas as publicações produzidas em território nacional, por qualquer meio ou processo, objetivando assegurar a coleta, a guarda e a difusão da produção intelectual brasileira, visando a preservação e formação da Coleção Memória Nacional. Estão inclusas obras de natureza bibliográfica e musical (BIBLIOTECA NACIONAL DO BRASIL).

Bettencourt (2014, p. 15) observa que o conceito alexandrino de “lugar de memória”, “templo do saber”, norteou os princípios fundadores das Bibliotecas Nacionais, instituições que surgiram a partir do século XVIII, como fatores de constituição da identidade nacional. A autora destaca, porém, que a função das Bibliotecas Nacionais, “como locais primordiais para a conservação, difusão e democratização de acervos históricos e culturais”, vem passando por alterações depois do advento da internet.

Face a essas mudanças, as Bibliotecas Nacionais ampliaram as suas missões e vêm procurando formas de se adequar aos novos tempos, buscando utilizar as novas tecnologias para promover sua missão primordial: preservar e assegurar o acesso à memória documental que se encontra sob sua guarda ao maior número possível de cidadãos (BETTENCOURT, 2014, p. 16).

A pesquisadora refere-se ao desenvolvimento de unidades digitais das Bibliotecas Nacionais, iniciativas que começaram a surgir nas últimas duas décadas. “As chamadas ‘bibliotecas digitais’, que passaram a disponibilizar fontes primárias e secundárias na internet, em “*Web*” sites e bases de dados *on-line*, o que foi possível por meio da digitalização e virtualização dos acervos” (BETTENCOURT, 2014, p. 16).

5.2 BIBLIOTECAS DIGITAIS

Biblioteca digital é um ambiente que integra coleções, serviços e pessoas na sustentação do ciclo complexo de criação, disseminação, uso e preservação de dados, informação e conhecimento, como resume Sayão (2009, p. 9), para logo em seguida concordar com Harter (1999) sobre as bibliotecas digitais serem uma ideia em movimento, ainda se desenvolvendo e tomando forma.

Pela definição de IFLA/UNESCO:

Uma biblioteca digital é uma coleção *on-line* de objetos digitais, de qualidade assegurada, que são criados, gerenciados e administrados de acordo com princípios aceitos internacionalmente para o desenvolvimento de coleções e acessíveis de maneira coerente e sustentável, apoiados pelos serviços necessários a permitir que usuários encontrem e explorem essas fontes (IFLA/UNESCO).¹¹

¹¹ IFLA/UNESCO *Manifesto for Digital Libraries*. Disponível em: <<http://www.ifla.org/publications/iflaunesco-manifesto-for-digital-libraries>>. Acesso em: 09 fev. 2016.

Bettencourt (2011, p. 41) afirma que as bibliotecas digitais foram concebidas e construídas somente a partir da década de 1990 e, embora recentes, constituem-se, juntamente com as tecnologias que as sustentam, o maior passo dado na história da humanidade para trazer à realidade o sonho do acesso universal ao conhecimento, preconizado por tantos visionários ao longo dos últimos séculos, como os belgas Paul Otlet e Henri La Fontaine, com seu *Mundaneum* no início do século XX, ou o norte-americano Vannevar Bush ao vislumbrar seu *memex* na metade da década de 1940.

Além de Otlet, La Fontaine e Bush, outro utopista é citado por Cunha (2008, p. 4): H. G. Wells, que em 1938 sugeriu a criação de uma enciclopédia universal contendo a memória planetária completa para toda a humanidade.

Para Cunha:

Uma biblioteca digital – uma coleção de informação digitalizada e organizada – tem um potencial informacional que dificilmente terá sido alcançado por alguma biblioteca convencional, isto é, ela pode entregar a informação diretamente na mesa do usuário (à semelhança da máquina *memex*, imaginada por Vannevar Bush), possui a capacidade de executar estratégias de busca por palavras isoladas ou por expressões inteiras, e o seu conteúdo informacional – seja ele na forma textual, sonora ou em imagens – não sofre os desgastes naturais decorrentes do uso intensivo do documento impresso (CUNHA, 2008, p. 5).

Sayão trata sobre a visão da comunidade da Biblioteconomia e da Ciência da Informação sobre as bibliotecas digitais. Esses grupos visualizam a biblioteca digital menos como um sistema de computação e mais como uma instituição, como uma extensão lógica do que as bibliotecas vêm fazendo, ou seja, adquirindo, organizando e disseminando conhecimento por meio do uso das tecnologias correntes.

“Para o bibliotecário, a biblioteca digital é um estágio a mais no desenvolvimento contínuo de novos meios de publicação, bem como uma nova infraestrutura tecnológica e organizacional voltada para disseminar informação e conhecimento” (SAYÃO, 2009, p. 10). Cunha (2008, p. 5) relembra que a biblioteca convencional é aquela em que a maioria dos itens do acervo apresenta-se em papel, apesar de esta também ter evoluído ao longo dos séculos passando a adotar recursos como catálogos e bancos de dados em linha. “A biblioteca digital combina a estrutura de coleta da informação, tradicionalmente usada pelas bibliotecas e arquivos, com o uso da representação digital tornada possível pela informática.”

Sayão (2009, p. 11) menciona ainda a questão “da ‘googlização’ das bibliotecas digitais”, ou seja, “a incômoda e errônea concepção de que o *Google* representa a apoteose da informação digital e os problemas existentes nesse domínio já foram resolvidos por esse serviço ou por outra ferramenta semelhante.” Nessa visão, a biblioteca digital aparece limitada às funções de busca e acesso, quando na verdade essas são apenas duas das

funcionalidades de uma biblioteca plena.

Em uma biblioteca, as fontes de informação adicionadas à coleção são cuidadosamente selecionadas, organizadas e descritas – classificadas, catalogadas, indexadas e resumidas. Sayão lembra que isso não ocorre com frequência nas coleções encontradas na “Web”. É claro, afirma ele, que ao fazer tal observação não se está desprezando o importante papel da “Web” na disseminação de conteúdo de qualidade para os usuários. Aliás, a “Web” é essencial para a realização de uma biblioteca digital. Só não se deve confundir duas ferramentas com características diversas.

Continua Sayão (2009, p. 13): “Apesar de a biblioteca digital ser, na maioria das vezes, um serviço vinculado à biblioteca tradicional, fica claro que existe uma distinção entre elas”. Para Sayão, a transição do impresso para o digital implica na criação de camadas de funcionalidades, de modos diferenciados de disseminação e entrega da informação e na forma como nos relacionamos com ela.

A passagem inicial do impresso para o digital teve como ênfase a conversão retrospectiva direta de conteúdos impressos para formatos digitais, por exemplo, a conversão de documentos raros, frágeis ou muito consultados. A versão digitalizada dos estoques informacionais da biblioteca tradicional proporcionou a possibilidade inédita do acesso independente de distância e de tempo, o compartilhamento por mais de um usuário de uma mesma obra a um custo muito baixo e, é claro, o acesso instantâneo e fácil a uma versão digital do texto completo (SAYÃO, 2009, p. 13).

Os jornais antigos digitalizados por diversas Bibliotecas Nacionais – objeto de estudo deste trabalho – servem de exemplo para o descrito pelo pesquisador, para quem as bibliotecas digitais incluem as funcionalidades das tradicionais, mas potencialmente vão além em escopo e em significado.

Para Sayão (2009, p. 14), o ambiente da biblioteca digital é um espaço dinâmico, constituído de informações eletrônicas, com níveis diferenciados de granularidade e serviços que possibilitam inúmeras configurações nas suas formas de disseminação e uma gama extraordinária de usos e reúsos para os seus estoques informacionais e para as representações correspondentes.

Nota-se a importância de levar essas características em consideração, fugindo do risco de transformar acervos digitalizados em meros depósitos de textos, nos quais a recuperação da informação torna-se uma tarefa impossível de ser realizada com qualidade.

5.2.1 Digitalização de acervos e o uso do OCR

Como apresentado nas *Recomendações para Digitalização de Documentos Arquivísticos Permanentes do Conselho Nacional de Arquivos* (Conarq, 2010), a digitalização

é dirigida ao acesso, difusão e preservação do acervo documental. Porém, o produto da conversão não será igual ao original e não substitui o original a ser preservado.

A digitalização torna-se um meio a mais de preservar os conteúdos raros, únicos ou frágeis ao mesmo tempo em que proporciona acesso universal a representações digitais desses conteúdos. O escaneamento de documentos possibilita que pessoas de qualquer lugar do mundo, com perfis e necessidades variadas de informação, possam navegar por todo o conteúdo oferecido nas bibliotecas disponíveis na “Web”, desde que tenham acesso a um computador e a uma conexão à internet.

Com esta possibilidade, torna-se mais fácil e rápido o acesso ao acervo, não limitando o usuário apenas aos endereços físicos das bibliotecas/arquivos. Conteúdos de bibliotecas locais ou regionais passam a ser acessados de qualquer lugar do mundo a qualquer momento.

Rodríguez Fuentes (2011, p. 24) afirma que quando uma grande biblioteca decide escanear suas coleções, o primeiro passo é definir qual parte do acervo poderá ser digitalizada com os recursos disponíveis, diante da impossibilidade de tratar das coleções completas. No caso da coleção administrada por Rodríguez Fuentes, dentro do acervo da Biblioteca Nacional da Espanha, dos oito mil títulos de imprensa histórica existentes, o volume de dinheiro disponível permitiu reproduzir 10% da coleção.

A partir da decisão sobre o tamanho da coleção a ser digitalizada, parte-se para a seleção. A lista de prioridades de digitalização varia de uma biblioteca para outra, mas de modo geral orienta-se pelo estado de conservação de determinado item, pelo tipo de material disponível, pela procura e uso do item no acervo e também por influencia de patrocinadores dispostos a contribuir para a digitalização de uma coleção específica.

Selecionados os exemplares, eles devem passar por uma preparação, seja para microfilmagem ou para digitalização. Documentos avulsos ou encadernados devem passar por tratamento de higienização, consolidação de suporte e planificação.

A higienização trata da eliminação mecânica de todas as sujidades que se encontram nos documentos e dos agentes considerados agressores, tais como os cliques oxidados ou não, os excrementos de insetos, os grampos metálicos, os elementos generalizados utilizados como marcadores de folhas, as poeiras, as partículas sólidas, e todos os elementos espúrios à estrutura física dos documentos. A consolidação do suporte consiste na recomposição das partes faltantes e na união dos rasgos e cortes, e a planificação consiste na correção de dobras e amassados (SPINELLI; BRANDÃO; FRANÇA, 2011, p. 15).

Financiamento é um dos desafios enfrentados pelas bibliotecas, assim como as questões de direitos autorais e a qualidade dos jornais antigos a serem digitalizados relacionada à conservação do papel e ao tipo de letra/fonte utilizada por publicações antigas.

Outro aspecto positivo a ser considerado é a oportunidade de integrar coleções de diferentes bibliotecas.

O armazenamento digital amplia as possibilidades de pontos de acesso a um determinado documento. Nos sistemas manuais tradicionais e mesmo nos catálogos automatizados produzidos até o final dos anos 1980, as descrições mínimas restringiam-se a dados sobre o autor, título e alguns cabeçalhos de assunto. Atualmente, dezenas de termos de indexação podem ser incluídos e, também, diversos níveis de representação do documento. Tais características agregam, sobremaneira, um alto grau de flexibilidade e qualidade na busca e recuperação da informação. Além disso, no caso de periódicos eletrônicos, por exemplo, as várias partes do documento, como objeto físico, podem ficar hospedadas em diversos computadores (CUNHA, 2008, p. 7).

Além disso, a digitalização, como será visto nas hemerotecas apresentadas nas páginas seguintes, impulsiona o uso de novos recursos. Exemplo disso é o uso da tecnologia de Reconhecimento Óptico de Caracteres (OCR, *Optical Character Recognition*), que possibilita realizar buscas no conteúdo de textos impressos, permitindo que grandes quantidades de informações presentes em textos digitalizados tornem-se recuperáveis.

O OCR funciona, em termos técnicos, como descrito pela *ABBYY FineReader OCR*:

Primeiro, o programa analisa a estrutura da imagem do documento. Ele divide a página em elementos, como blocos de texto, tabelas, imagens etc. As linhas são divididas em palavras e então em caracteres. Uma vez que os caracteres são individualizados, o programa os compara com um conjunto de imagens-padrão. São levantadas várias hipóteses sobre o que seria aquele caractere. Baseado nessas hipóteses, o programa analisa diferentes variantes de quebra de linhas em palavras e palavras em caracteres. Depois de processar um número enorme dessas hipóteses, o programa finalmente toma uma decisão, apresentando ao usuário o texto reconhecido¹² (ABBYY FINEREADER OCR, 2015).

O software do OCR combina funções dos olhos e cérebro humanos, como explica Holley (2008, p. 1-2), por isso pode ser considerado como um programa de inteligência artificial. Um indivíduo pode reconhecer nas páginas de jornal facilmente textos em diferentes fontes e com graus variados de qualidade de impressão. Assim, humanos reconhecem, traduzem e interpretam um texto rapidamente, mesmo que sejam textos veiculados em um jornal de 1800 com qualidade inferior de impressão, por exemplo. O OCR imita esta atividade, mas com um grau inferior de perfeição se comparado aos olhos e mente humanos.

Estima-se que o OCR tenha sido utilizado em digitalização de jornais antigos pela primeira vez no começo dos anos 1990. Profissionais da Biblioteca Nacional Britânica, em conjunto com bibliotecários australianos, realizaram um experimento em duas coleções de jornais (*Burney e Ferguson Collections*), mas não obtiveram bons resultados. As razões foram diversas, como qualidade do papel, condições do jornal ou microfilme original, curvatura dos

¹² Disponível em: <<http://www.abbyy.com/finereader/about-ocr/what-is-ocr/>>. Acesso em: 18 ago. 2015.

volumes encadernados, tamanho e estilo da fonte, impressão com mais ou menos tinta, leiaute das colunas. Porém, não se pode negar a utilidade do OCR.

Baseada na experiência de uso do OCR na Biblioteca Nacional da Austrália, Holley (2008, p. 2-3) enumerou os fatores que diminuem o desempenho do software e apontou sugestões e soluções (Quadro 6).

Quadro 6 – Desempenho do OCR, fatores e indicações

Etapas do processo	Fatores que influenciam o desempenho do OCR	Ações recomendadas em projetos com jornais antigos
Obter o original	Qualidade do original	<ul style="list-style-type: none"> • Usar as cópias impressas originais, se o orçamento permitir, pois a digitalização de cópias impressas é consideravelmente mais cara do que a de microfilme. • Cópias impressas a serem microfilmadas ou digitalizadas devem estar o mais completas e limpas possível – de preferência não encadernadas. • Usar microfilmes feitos após a criação e a utilização de padrões de imagem (após 1990). • Usar apenas a cópia negativa do microfilme (primeira geração) ou a versão impressa original. Nunca cópias.
Digitalizar a imagem	Resolução da digitalização e formato do arquivo	<ul style="list-style-type: none"> • Resolução da digitalização deve ter 300 ppi ou acima para capturar tantas informações da imagem quanto possíveis. • Escolher um formato de arquivo sem perdas, como TIFF, por exemplo, para não perder informações (<i>pixels</i>) da imagem.
Criar bom contraste entre branco e preto na imagem (Imagem pré-processada)	<ul style="list-style-type: none"> • Profundidade de bits da imagem. • Otimização da imagem / processo binário. • Qualidade da fonte (densidade do microfilme). 	<ul style="list-style-type: none"> • Escanear imagem em escala de cinza ou bitonal (preto e branco). • Otimização da imagem para OCR, a fim de aumentar contraste e densidade, realizar antes de transformar em OCR, tanto antes de escanear quanto antes de customizar o programa. • Se a imagem for convertida em escala de cinza, então converter a imagem em bitonal (binarização, preto e branco). • Obter a melhor qualidade da fonte. • Checar a densidade do microfilme antes de digitalizar.
OCR - Analisar e discriminar o leiaute da página.	<ul style="list-style-type: none"> • Página distorcida. • Página com leiaute complexo. • Espaços em branco entre linhas, colunas e margens da página para identificação dos limites. 	<ul style="list-style-type: none"> • Desamassar as páginas na preparação da imagem para que as palavras e frases fiquem na horizontal. • Leiaute das páginas e espaços em branco não podem ser alterados. Trabalhe com o que você tem.

OCR - Analisar a borda de cada caractere.	<ul style="list-style-type: none"> • Otimização da imagem. • Qualidade da fonte. 	<ul style="list-style-type: none"> • Prepare bem a imagem para ser vista pelo OCR, para que as bordas dos caracteres sejam suavizadas, arredondadas, bem definidas e com bom contraste. • Obtenha a melhor fonte possível (páginas amassadas, manchadas, mofadas, apagadas e fora de foco afetam diretamente na leitura dos caracteres pelo software).
OCR - Comparar limites dos caracteres para padronizar imagens e tomar decisões de qual caractere se trata.	<ul style="list-style-type: none"> • Padrões no banco de dados do software OCR. • Algoritmos no OCR. 	<ul style="list-style-type: none"> • Selecionar um bom programa OCR.
OCR - Comparar palavras inteiras às presentes em dicionários e tomar decisões sobre a confiabilidade.	<ul style="list-style-type: none"> • Algoritmos e dicionários construídos no OCR. 	<ul style="list-style-type: none"> • Selecionar um bom programa OCR.

Fonte: Holley (2008, p. 2-3).

Apesar de as empresas que fornecem o software exibirem altas taxas de precisão, sabe-se que esses valores são referentes a documentos de fácil leitura, como documentos em Word ou quando há alguma participação humana no processo – ainda não há melhor identificador de palavras e mais preciso do que o ser humano.

Quando se tratam de jornais, especialmente publicações antigas, o grau de fidelidade tende a ser mais baixo do que os 99% apregoados pelas empresas de OCR. Em teste realizado pela Biblioteca Nacional da Holanda, em uma coleção de jornais datados entre 1803-1954, o nível de precisão foi 68%, taxa semelhante à registrada em testes realizados pela Biblioteca Nacional da Austrália.

Como relatado anteriormente, a Biblioteca Nacional da Espanha chegou a 86% de precisão nos resultados. Apesar de ainda não ser a solução ideal, Rodríguez Fuentes (2011, p. 27) comenta que “esta é a única ferramenta de busca de conteúdo que podemos oferecer a nossos usuários”.

Uma das melhores soluções seria colocar “corretores humanos”, mas o consumo de tempo e de dinheiro não permite a adoção desta opção – talvez apenas se fosse para corrigir uma coleção específica e pequena. Alternativa encontrada pela Biblioteca Nacional da Austrália foi permitir que usuários sugiram correções. Tão logo a ideia foi colocada no ar em 2008, houve adesão do público. Em 12 semanas, 1,2 mil usuários haviam participado, corrigindo surpreendentemente 700 mil linhas dentro de 50 mil artigos. A biblioteca recebeu também várias mensagens do público.

“Os usuários não esperavam poder corrigir os textos do OCR. Assim que descobriram, rapidamente entenderam o conceito e o método. Alguns sentiram viciados e recompensados em corrigir os textos. Eles corrigiram muito mais do que poderíamos esperar”, avalia Holley (2008, p. 11). Usuários da biblioteca têm chance de atingir 100% de precisão por conhecer o idioma, a história e o contexto das publicações, diferente dos profissionais contratados pela empresa de OCR, que chegam em média a 99,5%.

Alternativa para oferecer maior valor agregado aos usuários é a descrição das publicações, textos curtos que façam uma apresentação do jornal ou revista, apontem sua importância, sua história e o papel que representou nas épocas em que foi publicado. Ou, ainda, inserção dos jornais no catálogo bibliográfico da unidade de informação a que pertencem.

5.3 HEMEROTECAS

O termo hemeroteca tem origem no vocabulário grego. Os vocábulos, *heméra* que tem o significado de dia somado à *théke* que significa “coleção” ou “depósito” criam a palavra hemeroteca, ou seja, um conjunto organizado ou coleção de periódicos.

As publicações abrangidas nas hemerotecas costumam ser de diferentes tipos:

Jornais e diários (entregam informações da atualidade, são uma fonte valiosa de informação sobre eventos, problemas, notícias ou ideias ao longo do tempo), publicações periódicas chamadas primárias, como trabalhos originais (revistas), publicações seriadas secundárias (recompilação de conteúdos das revistas primárias em forma de resumos) e publicações periódicas terciárias (recompilação de artigos, documentos, informes técnicos geralmente sobre um tema, como revisões de literatura) (KEYER ZULOAGA, 2009, p. 17).

Pesquisadores da área costumam diferenciar publicações seriadas de publicações periódicas. Seriada, de acordo com Cunha (2008, p. 303), é a publicação editada em fascículos sucessivos, contendo indicação numérica ou cronológica e, em regra, com a intenção de continuidade indefinida. Estão inseridas dentro da categoria de publicação seriadas as revistas especializadas, os jornais, os anuários (relatórios, balanços e almanaques), os periódicos gerais, as memórias, atas e relatórios de entidades coletivas. Podem ser incluídas nesta lista também as séries monográficas e de editores comerciais.

Keyer Zuloaga (2009, p. 17) elenca as características das publicações seriadas:

- a) Constituem-se um dos principais instrumentos de transferência de informação;
- b) São mais especializadas e atualizadas do que livros;
- c) São editadas normalmente em sequências cronológicas regulares;

- d) Caracterizam-se pela variedade de conteúdos e artigos;
- e) Contam com a participação de vários autores.

A descrição de publicação periódica refere-se à publicação seriada, veiculada em intervalos regulares, com periodicidade fixa inferior a um ano, como define Herrero Pascual (1996, p. 66). Publicações periódicas, de acordo com Keyer Zuloaga (2009, p. 17), são publicações coletivas que aparecem a intervalos regulares durante um tempo não limitado e com fascículos ou números que constituem um ou vários volumes em uma série contínua. Incluem-se nesta definição desde revistas científicas até jornais e revistas não especializados.

Na Ciência da Informação, a palavra periódico é comumente usada como sinônimo de publicação científica¹³. Neste trabalho serão usados os termos jornais e revistas quando em referência aos itens que compõem as hemerotecas digitais das Bibliotecas Nacionais.

Hemerotecas podem ser encontradas em bibliotecas – assim como nos arquivos de meios de comunicação. Geralmente são seções dedicadas a conservação, organização e consulta de materiais temáticos como revistas, jornais, folhetins, suplementos e anuários. Também abrigam recortes de textos impressos veiculados em diferentes formatos com o objetivo de facilitar pesquisas e trabalhos, especialmente em empresas, colégios e instituições.

Em termos gerais, as hemerotecas tem que cumprir as seguintes funções, como elencadas por Aparício Fernandez (1996, p. 97): formação e gestão da coleção; controle bibliográfico; conservação; indexação seletiva; difusão do acervo; atuação como centro de informação; oferecimento aos usuários da possibilidade de consulta a exemplares da imprensa nacional e estrangeira; e cooperação e intercâmbio de informação com outras hemerotecas e centros de documentação.

Importante ressaltar que uma das principais funções das hemerotecas é possibilitar o resgate de momentos históricos importantes e a preservação da memória de povos e culturas. A ferramenta tem grande utilidade para escritores, historiadores, pesquisadores, jornalistas, entre outros profissionais que necessitam deste tipo de informação.

As hemerotecas pertencentes a Bibliotecas Nacionais seguem as mesmas diretrizes aplicadas aos demais acervos, ou seja, têm como objetivo prioritário a reunião, conservação e difusão de toda a produção de publicações – neste caso, periódicas – do país. Jornais e revistas atendem à lei do depósito legal. Dentro das possibilidades financeiras de cada

¹³ Do *Dicionário de Biblioteconomia e Arquivologia*: “p. científica *scientific publication* documento que aborda um assunto de interesse científico, utilizando metodologia científica, redigido segundo as convenções usuais de sua disciplina, e que é distribuído à comunidade da área. Sinônimos: artigo científico, comunicação científica, comunidade científica, informação científica” (CUNHA; CAVALCANTI, 2008, p. 302).

Biblioteca Nacional, também se mantém um acervo de publicações periódicas estrangeiras – apesar de a digitalização e a possibilidade de acesso a conteúdos digitais de maneira *on-line* terem alterado esta prática.

Guallar, Abadal e Codina, pesquisadores espanhóis autores de artigos sobre coleções de jornais, escrevem que:

Os sistemas de acesso e consulta aos conteúdos retrospectivos dos jornais na internet são conhecidos indistintamente como arquivos ou hemerotecas digitais. Desde as suas origens transformaram-se de meros depósitos de documentos sem sistema de recuperação (e que permitiam apenas a navegação cronológica) a uma importante variedade de formas de recuperar informação, com ampla cobertura temporal e diferentes possibilidades de visualização de conteúdos (GUALLAR; ABADAL; CODINA, 2012, p. 596, tradução nossa).

De acordo com Lapoglia, uma hemeroteca permite a socialização dos dizeres do passado, instigando sujeitos a repensar a história, buscar o contexto em que os escritos foram produzidos, comparar o passado com o presente e observar o que mudou e o que continua nos discursos da/sobre a história.

Lapoglia afirma serem jornais e revistas:

[...] fonte documental afetada pelo ponto de vista dos contemporâneos da época, construída sem as intervenções (ou reflexões) dos acontecimentos subsequentes, ou seja, diferentes dos livros de história ou dos relatos de sujeitos que viveram a época, os dizeres dos periódicos conservam as condições de produção do momento em que foi escrito, sem a influência dos acontecimentos posteriores e das reflexões oriundas deles (LAPOGLIA, 2012, p. 126).

Alonso Navarro (1996, p. 37), ao analisar o valor histórico das hemerotecas, afirma que ter contato com uma hemeroteca, consultar velhas coleções de diários, semanários ou outras publicações periódicas, trazem para mais perto as vivências daqueles que deixaram seu comentário, seu juízo crítico, suas paixões e ilusões registradas para as gerações seguintes.

Ao ler jornais e revistas antigos – considerando as subjetividades de cada época – entra-se em um mundo de grandezas e misérias, acertos e desgraças, metas alcançadas e frustrações, difíceis de medir sem a ótica da distância, a qual permite analisar de forma tranquila o valor histórico do que se está lendo. As passagens registradas nos jornais permitem conhecer melhor vários aspectos da história humana, pequenos retratos do cotidiano, momentos especiais de pessoas comuns.

Curiel Defossé e Jiménez Rivera, da Hemeroteca Digital da Biblioteca Nacional do México, complementam esse pensamento ao escrever que o acesso a jornais e revistas representa um desafio aos responsáveis da difusão de seus conteúdos devido a amplitude temática disposta em uma estrutura e diagramação complicadas e variáveis através dos tempos.

A imprensa aborda todos os assuntos que caracterizaram e formaram a sociedade, com diversos pontos de vista e em todas as formas possíveis de comunicação impressa, textual e gráfica. Pelos dizeres de Vicente Quirarte, “o melhor instrumento para tomar o pulso da sociedade, a pegada que o passar dos dias imprime nos seus leitores. A letra impressa que modifica ações, orienta e forma opinião, constrói sua própria história” (CURIEL DEFOSSÉ; JIMÉNEZ RIVERA, 2012. p. 1).

No capítulo a seguir são apresentados exemplos de hemerotecas criadas e mantidas por Bibliotecas Nacionais de diferentes países, selecionados segundo os critérios da presente pesquisa e analisados segundo suas características.

6. BIBLIOTECAS NACIONAIS COM HEMEROTECAS DIGITAIS

As hemerotecas digitais de Bibliotecas Nacionais são o objeto de estudo desta tese. Por meio da apresentação e análise de 13 delas pretende-se encontrar subsídios para apontar alcances e limites desta unidade de informação como fonte para cientistas.

A fim de detectar quais Bibliotecas Nacionais oferecem acesso a coleções digitalizadas de jornais, foi realizado um levantamento incluindo os 193 países listados no site da Organização das Nações Unidas (ONU)¹⁴. As buscas foram realizadas no *Google*, a partir da inserção do nome do país mais o termo *national library* escrito entre aspas. Exemplo: Austrália + “*national library*”.

Nas buscas, realizadas entre os dias 14 e 16 de agosto de 2015, foram obtidos os resultados apresentados a seguir:

a) 13 sites de Bibliotecas Nacionais com acesso a hemerotecas digitais, como listados no Quadro 7.

b) 10 sites de Bibliotecas Nacionais com bibliotecas digitais, mas sem acesso a coleções digitalizadas de jornais – Canadá, Chile, Colômbia, Cuba, Itália, Jamaica, Panamá, Portugal, Trinidad e Tobago e Uruguai.

c) 85 sites de Bibliotecas Nacionais agrupados por apresentarem a página inicial em idioma local de difícil compreensão ou a página inicial em inglês, mas com conteúdo da biblioteca digital ou equivalente em idioma local, como finlandês, holandês ou norueguês, ou ainda por oferecerem apenas informações de itens catalogados e disponíveis no acervo físico ou somente informações sobre o atendimento realizado nas unidades presenciais;

d) 85 casos em que os sites das Bibliotecas Nacionais não foram recuperados na busca ou que as informações sobre as Bibliotecas Nacionais não eram apresentadas em páginas próprias, mas nos sites do Ministério da Cultura ou por meio de uma *fan page* no *Facebook*.

Em países como Afeganistão e Bósnia e Herzegovina, as Bibliotecas Nacionais foram destruídas em guerras ocorridas nos últimos 20 anos. Apesar da reconstrução recente, ainda não têm página na “Web”.

A partir desses resultados, foram coletados dados sobre as 13 Bibliotecas Nacionais (Quadro 7) que mantêm hemerotecas digitais. No capítulo 5 são apresentadas as hemerotecas de BN estrangeiras. No capítulo 6, a Hemeroteca Digital Brasileira.

¹⁴ Disponível em: <<http://www.un.org/en/members/>>. Acesso em: 14 ago. 2015

Quadro 7 – Bibliotecas Nacionais com hemerotecas digitais

País	Site BN	Hemeroteca	Site BN Digital / Hemeroteca
Argentina	http://www.bn.gov.ar	<i>Trapalanda</i>	http://trapalanda.bn.gov.ar/jspui/handle/123456789/5
Austrália	http://www.nla.gov.au	<i>Trove</i>	http://trove.nla.gov.au/
Áustria	http://www.onb.ac.at/	<i>ANNO</i>	http://anno.onb.ac.at/
Bélgica	http://www.kbr.be	<i>BelgicaPress</i>	http://belgica.kbr.be/fr/coll/jour/jour_fr.html
Brasil	http://www.bn.br	<i>Hemeroteca Digital</i>	http://bndigital.bn.br/hemeroteca-digital/
Costa Rica	http://www.sinabi.go.cr/bibliotecas/biblioteca_nacional.aspx#.UzDWXfldUg8		http://www.sinabi.go.cr/contextual/busqueda.aspx?tipobuscador=4
Espanha	http://www.bne.es	<i>Hemeroteca Digital</i>	http://hemerotecadigital.bne.es
Estados Unidos da América	http://loc.gov	<i>Chronicling America</i>	http://chroniclingamerica.loc.gov
França	http://www.bnf.fr	<i>Gallica</i>	http://gallica.bnf.fr/html/presse-et-revues/presse-et-revues
Liechtenstein	http://www.eliechtensteinensia.li	<i>Zeitungen</i>	http://www.eliechtensteinensia.li/Zeitungen/
México	http://bnm.unam.mx	<i>Hemeroteca Nacional Digital</i>	http://www.hndm.unam.mx
Nova Zelândia	http://natlib.govt.nz	<i>Papers past</i>	http://paperspast.natlib.govt.nz
Reino Unido	http://www.bl.uk	<i>British Newspaper Archive</i>	http://www.britishnewspaperarchive.co.uk/

6.1 ARGENTINA: TRAPALANDA

Biblioteca Nacional da Argentina: <http://www.bn.gov.ar/>

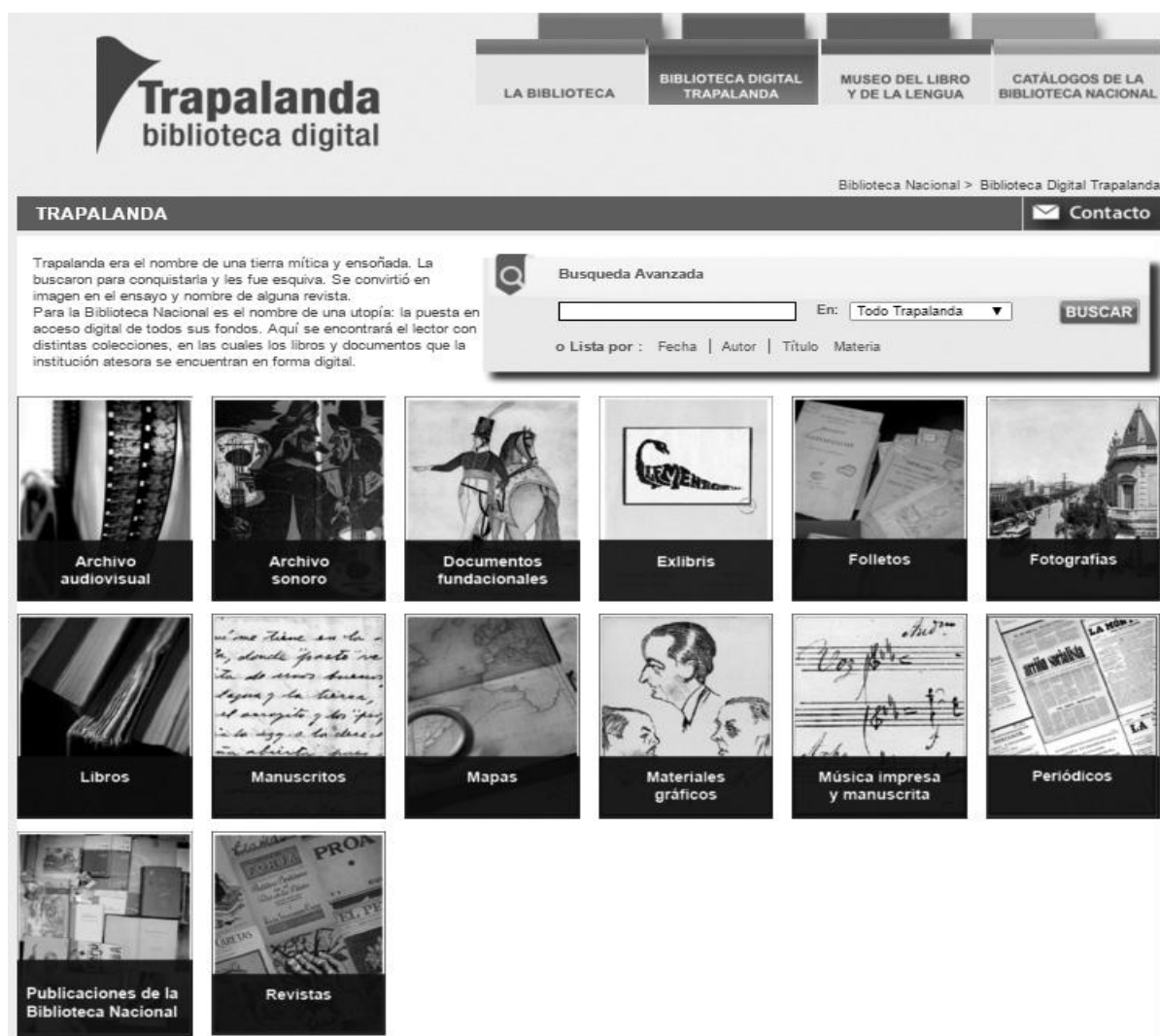
Trapalanda: <http://trapalanda.bn.gov.ar/jspui/handle/123456789/1>

Trapalanda é o nome de uma terra mítica e de sonho localizada ao sul da Argentina, na região da Patagônia. Trapalanda também foi o nome escolhido pela Biblioteca Nacional da Argentina para batizar a sua biblioteca digital. Como explica em seu site, Trapalanda também é o nome de uma utopia da BN: “colocar em acesso digital todas as suas coleções”.

A página da biblioteca digital foi lançada em agosto de 2014 com seis mil obras disponíveis *on-line*. O processo de digitalização foi iniciado em 2011, mas as informações do acervo começaram a ser organizadas no ano seguinte. Antes da criação da Trapalanda, a Biblioteca Nacional da Argentina mantinha uma biblioteca virtual com aproximadamente mil registros. A administração da biblioteca optou por usar o software livre *Dspace*, um sistema para construção de repositórios institucionais digitais.

A Trapalanda é composta por 14 seções, que incluem livros, manuscritos, mapas, folhetos, fotografias, arquivos sonoros e audiovisuais (Figura 12). As publicações periódicas foram distribuídas em duas seções: *Revistas* e *Periódicos*, sendo esta última a que abriga jornais.

Figura 12 – Página inicial da Trapalanda, biblioteca digital argentina



Fonte: Biblioteca Nacional da Argentina. Disponível em: <<http://goo.gl/V4j3Gu>>. Acesso em: 14 ago. 2015.

Ao clicar na seção *Periódicos*, o usuário pode ler que o acervo completo é composto por jornais veiculados desde a época da Argentina colônia até quase os dias atuais, edições europeias e americanas do século XVIII, publicações de associações, paróquias, partidos políticos, além de coleções particulares. Digitalizadas há edições que vão de 1801 a 1957. Na parte inferior da página, há ícones que remetem a dois períodos da coleção disponível *on-line*, o século XIX e o século XX. Ao clicar em um dos dois ícones, o leitor é direcionado para as capas dos jornais.

O mecanismo de busca está disponível tanto na página inicial da Trapalanda (no qual o usuário pode pesquisar em todo o acervo ou escolher a seção desejada) quanto em uma seção em particular (Figura 13). Apesar de se chamar busca avançada, a consulta é realizada por meio de palavra-chave, que recupera não o conteúdo, mas os títulos dos jornais.

Figura 13 – Página de busca da Trapalanda

PERIÓDICOS Contacto

La Biblioteca Nacional posee la más amplia colección de periódicos y diarios producidos en el país desde la época de la colonia hasta nuestros días; también conserva una muy amplia colección de periódicos europeos y americanos que van desde el siglo XVIII.

Todo tipo de periódico es conservado en la Biblioteca, los hay nacionales, provinciales y locales, también hay colecciones particulares y los de de interés general y aquellos que responden a intereses sectoriales como gremios, colectividades, parroquias, partidos políticos, asociaciones diversas o de determinado interés cultural o tema.

Búsqueda avanzada

En: PERIÓDICOS

o Lista por : Fecha | Autor | Título. Materia BUSCAR

Fonte: Biblioteca Nacional da Argentina. Disponível em: <<http://goo.gl/V4j3Gu>>. Acesso em: 14 ago. 2015.

Logo abaixo do campo de busca, há a opção de acesso a todos os itens digitalizados listados por data, título (nome da publicação), autor (nomes dos editores) ou assunto. As listas são apresentadas em até 100 itens, em ordem alfabética ou cronológica ascendente e descendente (Figura 14).

Figura 14 – Opção de listar o acervo por autor, data, local e assunto

BUSCAR "PERIÓDICOS" POR AUTOR

Ordenar por: En orden: Res. por página Autor/Registro:

Ir a: 0-9 A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

FECHA	TÍTULO	AUTOR(ES)
8-feb-1817	La Crónica Argentina. No. 40	Pazos Silva, Vicente, director; Valdés, Antonio José, redactor; Enriquez, Camilo, redactor
1-feb-1817	La Crónica Argentina. No. 39	Pazos Silva, Vicente, director; Valdés, Antonio José, redactor; Enriquez, Camilo, redactor
25-ene-1816	La Crónica Argentina. No. 38	Pazos Silva, Vicente, director; Valdés, Antonio José, redactor; Enriquez, Camilo, redactor
18-ene-1817	La Crónica Argentina. No. 37	Pazos Silva, Vicente, director; Valdés, Antonio José, redactor; Enriquez, Camilo, redactor
11-ene-1817	La Crónica Argentina. No. 36	Pazos Silva, Vicente, director; Valdés, Antonio José, redactor; Enriquez, Camilo, redactor
4-ene-1817	La Crónica Argentina. No. 35	Pazos Silva, Vicente, director; Valdés, Antonio José, redactor; Enriquez, Camilo, redactor
28-dic-1816	La Crónica Argentina. No. 34	Pazos Silva, Vicente, director; Valdés, Antonio José, redactor; Enriquez, Camilo, redactor
25-dic-1816	La Crónica Argentina. No. 33	Pazos Silva, Vicente, director; Valdés, Antonio José, redactor; Enriquez, Camilo, redactor
21-dic-1816	La Crónica Argentina. No. 32	Pazos Silva, Vicente, director; Valdés, Antonio José, redactor; Enriquez, Camilo, redactor

Fonte: Biblioteca Nacional da Argentina. Disponível em: <<http://goo.gl/V4j3Gu>>. Acesso em: 14 ago. 2015.

Figura 15 – Página de resultados de busca na seção *Periódicos* da Trapalanda

RESULTADOS DE BÚSQUEDA

Res. por página | Ordenar por En orden Autor/registro

Buscar: por

FECHA	TÍTULO	AUTOR(ES)
31-ene-1812	Gazeta de Buenos Ayres. No. 22	-
21-feb-1812	Gazeta de Buenos Ayres. No. 25	-
26-nov-1811	Gazeta de Buenos Ayres. No. 7	-
15-feb-1812	Gazeta de Buenos Ayres. Extraordinaria	-
3-sep-1811	Gazeta de Buenos Ayres. Extraordinaria	-
8-feb-1812	Gazeta de Buenos Ayres. Extraordinaria	-
1-ene-1813	Gazeta de Buenos Ayres. No. 39	-
28-feb-1812	Gazeta de Buenos Ayres. No. 26	-
17-ene-1812	Gazeta de Buenos Ayres. No. 20	-

Fonte: Biblioteca Nacional da Argentina. Disponível em: <<http://goo.gl/V4j3Gu>>. Acesso em: 14 ago. 2015.

A busca leva a uma lista de resultados com *links* que direcionam o visitante ao PDF das páginas de seu interesse (Figura 15). Abaixo da imagem (Figura 16), há informações sobre o item recuperado, como título, palavras-chave, data de publicação, editor, o endereço eletrônico em que se encontra o item, o registro completo no catálogo da BN (Dublin Core) e ainda a que coleções está ligado. Há o nome do arquivo salvo (ex: *NOz_1810_Real_Comunicada.PDF*), o tamanho (ex.: 843,4KB) e o formato (ex. Adobe PDF), além do botão para realizar o download da página recuperada.

Figura 16 – Página recuperada na Trapalanda

Num. 29. Pág. 113

GAZETA DE BUENOS=AYRES.

VIERNES 20 DE MARZO DE 1812.

*Rara temporum felicitate, ubi sentire quae velis,
et quae sentias, dicere licet.*

Tacito lib. 1. Hist.

CONTINUAN LAS OBSERVACIONES DIDACTICAS.

El éxito de nuestras armas, la disciplina militar, la administración interior, la opinión pública, la energía y el orden todo está íntimamente unido a las deliberaciones de la próxima asamblea. El pueblo la espera con un deseo íntimo; y si su esperanza puede ser un principio de cálculo, yo diría, que va a empezar una nueva serie de acontecimientos felices: yo diría que la victoria nos llama, y que los ejércitos están ya sobre el vestíbulo de su templo: yo diría que el espíritu público vuelve a su turno, y que la patria al fin va a sentarse sobre el trono que ocupaban los déspotas. Por el contrario, si no mejora mas conforme a los ritos convencionales que la política sanciona muchas veces como principios de equidad natural. La asamblea debe resolver y adoptar todas las medidas que puedan salvar la patria, sin temor de violar los derechos de los pueblos, cuya primera y última voluntad es conservar su existencia. Esta es la ley constitucional que rige en todas sus deliberaciones, y en virtud de ella queda autorizada para obrar según el imperio de las circunstancias y la urgencia de los peligros. Pero siendo estos tan palpables, es muy escandalosa la suspensión acordada, á pretexto de que el 23 que debía abrirse

< OTROS NÚMEROS >

INFORMACIÓN
 Título : Gazeta de Buenos Ayres. No. 29
 Otros títulos : Gaceta de Buenos AiresGaceta de Buenos AyresGazeta
 Ministerial del Gobierno de Buenos Aires
 Palabras clave :
[Dianios argentinos - Siglo XIX](#)
 Fecha de publicación : 20-mar-1812
 Editor : Imprenta de Niños Expósitos
 URI : <http://trapalanda.bn.gov.ar:8080/ispu/handle/123456789/1762>
 Aparece en las colecciones: 1812

Mostrar el registro Dublin Core completo del ítem

Archivo : BNA_MF00008_1812-03-20.pdf
 Tamaño : 1,43 MB
 Formato : Adobe PDF
 Visualizar la licencia **DESCARGAR**

Copyright Biblioteca Nacional Mariano Moreno de la República Argentina.

Fonte: Biblioteca Nacional da Argentina. Disponível em: <<http://goo.gl/V4j3Gu>>. Acesso em: 14 ago. 2015.

6.2 ÁUSTRIA: ANNO

Biblioteca Nacional da Áustria: <http://www.onb.ac.at/>

ANNO: <http://anno.onb.ac.at>

ANNO significa *AustriaN Newspapers Online*. É o nome da hemeroteca da Biblioteca Nacional da Áustria. “Trata-se do salão virtual de leitura de jornais e revistas austríacas, local onde se pode folhear e ler publicações *on-line*”, conforme definição apresentada na página inicial da hemeroteca (Figura 17).

Criado em agosto de 2003, o ANNO oferece acesso a publicações locais e regionais, compostas por textos sobre história, política, cultura e sociedade. Os organizadores da hemeroteca, como descrito no site, acreditam que não há opção melhor para se informar sobre o seu pedaço de mundo do que as notícias veiculadas nos jornais da sua própria região.

O processo de digitalização começou a ser realizado em maio de 2003, com o uso de *scanners* planetários das companhias Zeutschel, Microbox e Quidenus. Inicialmente o projeto previa a gravação do material digitalizado em resolução de 300 dpi, preto/branco e compressão sem perda em TIFF. À medida que foram realizadas as primeiras digitalizações, passou-se a gravar os documentos em formato colorido e em JPG, pois notou-se que não haveria alteração no custo (em relação ao preto/branco) e na qualidade, conforme explica Christa Müller (2015), responsável pelo ANNO e chefe da seção de serviços digitais da Biblioteca Nacional da Áustria.

De maio de 2003 a setembro de 2015 já foram digitalizados mais de 600 títulos, alcançando número superior a 14 milhões de páginas. O período coberto vai de 1568 a 1944, mas estão *on-line* títulos publicados entre 1689 e 1918 e entre 1938 e 1944. O período entre 1919 e 1937 ainda aguarda ser digitalizado. Os critérios para digitalização são: os jornais e revistas em condições físicas ruins; as publicações mais consultadas pelos usuários, a fim de evitar o manuseio e a possível deterioração do objeto; as edições relacionadas a projetos temáticos; as publicações de projetos de cooperação com outras organizações.

Anualmente tem sido digitalizadas 1,5 milhão de páginas. Christa Müller (2015) não consegue prever quanto tempo ainda levará para digitalizar todo o acervo de periódicos da BN da Áustria, mas especula que serão necessários muitos anos para que tudo esteja disponível *on-line*. A digitalização é realizada fora da biblioteca por uma empresa prestadora do serviço – a EMD¹⁵.

¹⁵ EMD é a empresa responsável pela digitalização do acervo da BN da Áustria: <<http://www.emd.at/>>. Acesso em: 3 out. 2015.

Figura 17 – Página inicial do ANNO

Österreichische Nationalbibliothek

Über ANNO

- Was ist ANNO?
- 10 Jahre ANNO
- FAQs
- Suchen in ANNO
- Drucken aus ANNO
- Kooperationspartner
- Kontakt
- Impressum

Service

- ANNOdazumal
- Kurzbeiträge zu historischen Zeitungen
- Projektteilnahmen von ANNO
- Zeitungen und Zeitschriften an der ÖNB
- Bestellung von Jubiläums- und Geburtstagszeitungen
- Weitere
- Zeitungsdigitalisierungsprojekte

ANNO
Historische Zeitungen
und Zeitschriften



ANNO - AustriaN Newspapers Online
Historische österreichische Zeitungen und Zeitschriften online

Listen

- Alphabetische Liste der Zeitungen und Zeitschriften
- Jahresübersicht der Zeitungen (1568-1944)
- **neu** Thematischer Einstieg

Suchen

- **neu** ANNO-Suche: Volltextsuche in historischen Zeitungen (1689-1918, 1938-1944)
- Suche in den Inhaltsverzeichnissen von einigen Zeitschriften

ANNO ist der virtuelle Zeitungslesesaal der Österreichischen Nationalbibliothek. Hier kann in historischen österreichischen Zeitungen und Zeitschriften online geblättert und gelesen werden.

Alle hier vorgestellten Beiträge finden Sie auch unter "ANNOdazumal" beziehungsweise "Kurzbeiträge zu historischen Zeitungen" in der linken Menüleiste.

2. Oktober 1850

Todestag von Joseph Madersperger, österreichischer Schneidermeister und Erfinder



Die *Wiener Zeitung* meldete am 8. Oktober 1850 den Tod von Joseph Madersperger, "bürgerlichen] Schneider und Bürgerfonds-Pfändner" (*Wiener Zeitung*, 8. Oktober 1850, S. 8). Der Erfinder der ersten "österreichische[n] Nähmaschine" (vgl. *Der Lehrerinnen-Wart*, 10. August 1894, S. 7) verstarb im Alter von 83 Jahren (vgl. *Wiener Zeitung*, 8. Oktober 1850, S. 8).

Madersperger erhielt das Privilegium für seine Erfindung bereits 1814 (vgl. *Grazer Mittags-Zeitung*, 4. Juli 1917, S. 3). Da diese aber noch nicht alle Erwartungen der zeitgenössischen Fachleute erfüllte (vgl. *Der Siebenbürger Bote*, 17. September 1853, S. 4), perfektionierte er sie und schloss die Entwicklung erst 1839 ab (vgl. *Österreichische Nähmaschinen- und Fahrrad-Zeitung*, 28. Februar 1918, S. 4).

Die gesellschaftliche Anerkennung von Maderspergers Erfindung wird in verschiedenen Zeitungen ganz unterschiedlich beschrieben. Manche bestehen darauf, dass die Nähmaschine per se von dem Österreicher erfunden wurde (z.B. *Journal des Österreichischen Lloyd*, 8. August 1840, S. 2 oder auch *Kikeriki*, 26. September 1880, S. 2). Andere weisen aber auch darauf hin, dass erst eine amerikanische Nähmaschine die von so vielen anderen Erfindern erhobene Wertschätzung des Publikums einholen konnte (vgl. *Der Siebenbürger Bote*, 17. September 1853, S. 4).

Es ist interessant zu sehen, wie das Schicksal Maderspergers und seiner Nähmaschine in den unterschiedlichen (politischen) Epochen beschrieben wurde. 1941 wird in der *Znaimer Tagblatt* Maderspergers "einigermaßen brauchbare" Nähmaschine der "kaum verwendungsfähiger" englischen Erfindung gegenübergestellt (*Znaimer Tagblatt*, 19. November 1941, S. 4). Die *Österreichische Nähmaschinen- und Fahrrad-Zeitung* merkt an, dass die beiden Nähmaschinen wegen ihrer Zweckwidmung kaum miteinander zu vergleichen wären. Die österreichische Variante war nämlich für das Nähen von Kleidungsstücken, die englische fürs Nähen von Schuhsohlen konzipiert (vgl. *Österreichische Nähmaschinen- und Fahrrad-Zeitung*, 30. Januar 1909, S. 29). Der wirklich von Erfolg gekrönte Nähmaschinentyp kam letztendlich aus Übersee:

1910 berichtet *Dillinger's Reisezeitung* über den Tod des Erfinders "der ersten brauchbaren Nähmaschine" (*Dillinger's Reisezeitung*, 1. September 1910, S. 7) in New York. Hiemit war Josef Thomas, dessen Erfindung durch die Singer-Gesellschaft die Weltberühmtheit errang, gemeint. Ebenda wird dennoch kurz angemerkt, dass die Entdeckung des Grundprinzips der Nähmaschine Madersperger zugeschrieben werden muss (vgl. ebda.).

Madersperger überließ seine Nähmaschine dem Wiener polytechnischen Institut, da es ihm offenbar nicht bewusst war, wie groß die Nachfrage nach einer solchen bahnbrechenden Erfindung war (vgl. *Morgen-Post* 21. September 1880, S. 3). Er selbst starb im Armenhaus von St. Marx (vgl. *Znaimer Tagblatt* 19. November 1941, Seite 4). Um seinen Namen zu verewigen, wurde ihm in seiner Heimatstadt Kufstein ein Denkmal errichtet (vgl. *Illustriertes Wiener Extrablatt*, 28. April 1903, S. 11).

Neu bei ANNO

- 13 neue oder ergänzte Titel online! (30.09.2015)
- 8 neue oder ergänzte Titel online! (31.08.2015)
- 37 neue oder ergänzte Titel online! (31.07.2015)
- Lesen Sie mehr...

Aktuelles

- 14 Millionen Seiten online! (26.01.15)
- 13 Millionen Seiten online! (03.10.14)
- 12,5 Millionen Seiten online! (04.08.14)
- 1914-1918 sowie Zeitschriften in der ANNO-Suche freigeschaltet! (27.05.14)

ANNO Startseite · Über uns · Kontakt · Impressum
Copyright © 2011 Österreichische Nationalbibliothek

Tanto a coleção física quanto a coleção digitalizada da BN austríaca são consultadas por diferentes públicos, estando entre eles pesquisadores científicos. Diariamente 2,5 mil leitores acessam o acervo digitalizado, realizando aproximadamente oito mil buscas. Ao acessar o site do ANNO, os usuários podem navegar pelos jornais e revistas em listas das publicações dispostas em ordem alfabética, por ano ou por assuntos.

A catalogação do acervo é realizada por meio da menção de nome do jornal (título), local e ano de publicação, idioma e assunto geral¹⁶ do veículo de comunicação (como, por exemplo, arquitetura, carro, educação, esperanto, história, saúde, guerra, literatura, agricultura ou viagem).

O formato adotado pelos catalogadores da Biblioteca Nacional da Áustria ainda é o MAB¹⁷ (*Maschinelles Austauschformat für Bibliotheken* – formato automático de intercâmbio para bibliotecas), desenvolvido na Alemanha com inspiração no MARC. De acordo com a Biblioteca Nacional alemã, o MAB foi “congelado” em 2006, não sendo mais possível realizar mudanças ou adições. Na Alemanha, o formato foi substituído pelo MARC21. Christa Müller (2015) confirma que o MAB é usado para a catalogação de todos os títulos de livros e jornais do acervo da biblioteca.

Para auxiliar as buscas, utiliza-se OCR. Por meios deste recurso os usuários conseguem realizar consultas na busca básica, por meio de palavras-chave (Figura 18), ou busca avançada (Figura 19), por meio de palavra-chave, título, idioma, data, local e assunto.

Figura 18 – Busca básica

Fonte: Biblioteca Nacional da Áustria. Disponível em: <<http://anno.onb.ac.at/>>. Acesso em: 15 ago. 2015.

Figura 19 – Busca avançada

Fonte: Biblioteca Nacional da Áustria. Disponível em: <<http://anno.onb.ac.at/>>. Acesso em: 15 ago. 2015.

¹⁶ Lista completa de assuntos catalogados. Disponível em: <http://anno.onb.ac.at/them_einstieg.htm>. Acesso em: 2 out. 2015.

¹⁷ Informações sobre o MAB no site da Biblioteca Nacional alemã. Disponível em: <http://www.dnb.de/EN/Standardisierung/Formate/MAB/mab_node.html>. Acesso em: 2 out. 2015.

Os resultados da busca são apresentados em lista (Figura 20), podendo ser ordenados posteriormente por relevância, data e ordem alfabética (dos nomes das publicações). No menu à esquerda, o usuário encontra diversos filtros, com os quais pode refinar a busca pelo tipo de veículo (jornal ou revista), título, local, idioma, período e assunto.

Figura 20 – Resultado da busca apresentado em forma de lista

The screenshot displays the ANNO search interface. At the top left, the search criteria are: Text: "Kaiser Karl Front" ~10, Title: (empty), Language: Alle, Issue type: (empty), and Date: (empty). The search results are sorted by Relevanz (Relevance). The interface includes a search bar, filters, and a list of search results with associated newspaper thumbnails.

Suchergebnisse 1 bis 10 von 211

Ergebnisse filtern: 1 2 3 >>>

Medium
Zeitung (211)
Zeitschrift (0)

Titel
(Neuglebe) Welt-Blatt (15)
Ausung aus der Tagespresse (13)
Wiener Bilder (11)
(Linzer) Tages-Post (10)
Vorarlberger Volksblatt (10)
Reichspost (9)
Neue Freie Presse (8)
Prager Tagblatt (8)
Teplitz-Schönauer Anzeiger (8)
Freunden-Blatt (7)

Ercheinungsart
Wien (101)
Linz (17)
Bregenz (15)
Budapest (12)
Praha (8)
Teplitz-Sanov (8)
Pest (7)
Salzburg (7)
Celje (6)
o.o. (6)

Zeitraum
1914 (1)

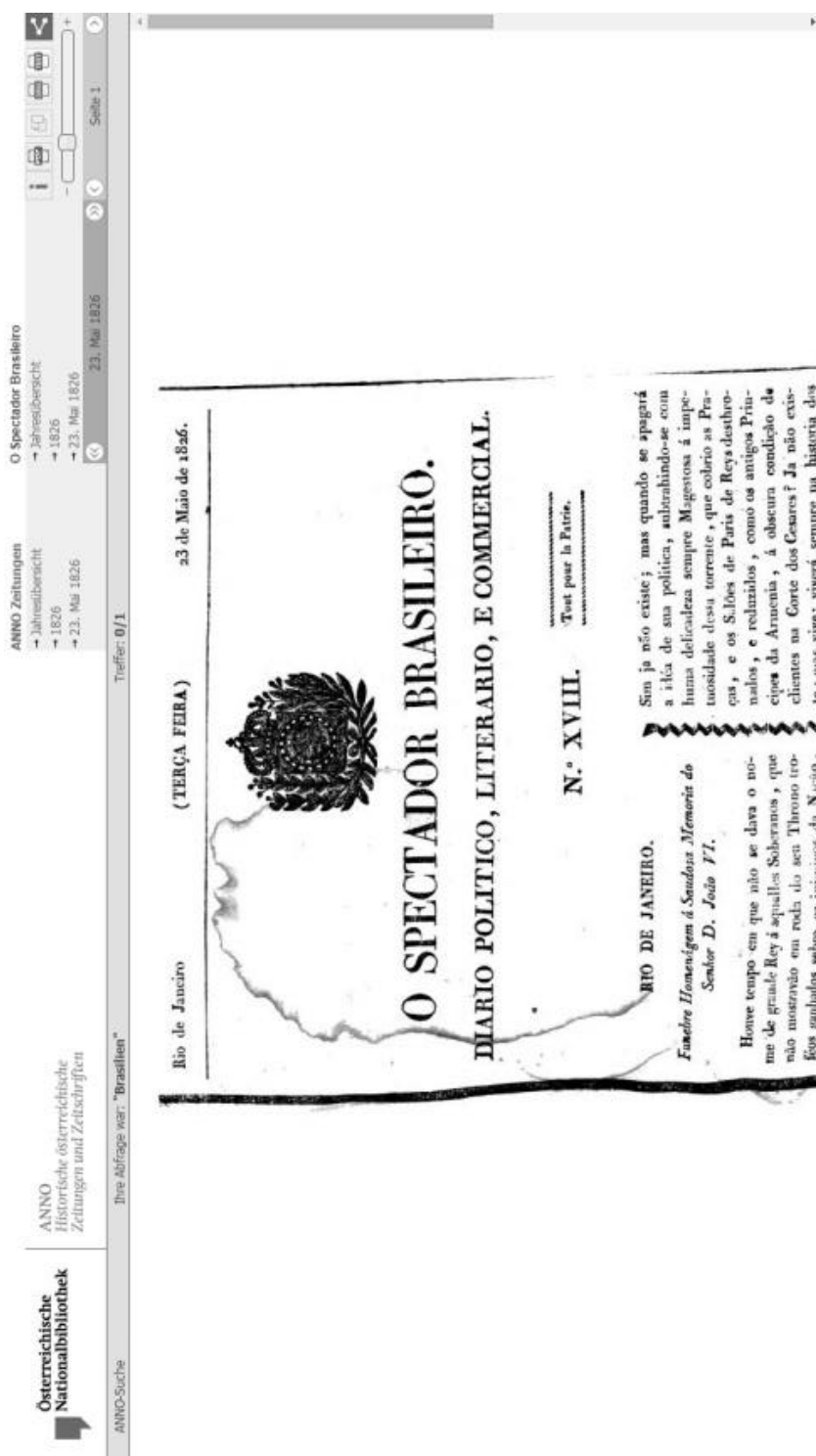
1
Neue Werte am Inn 30. Dezember 1917
 Zeitung
 Braunau am Inn - Deutsch - 2 Seiten
Thema: Tageszeitung
 2 Treffer im Text
 ...Unser **Kaiser** am der Tiroler **Front**
Kaiser Karl belobt einen mehrfach dekorierten Begünstigter.
 Lichtblicke des eheer. Kriegesquartiers...

2
Extrahaftungen des 1. Weltkrieges 11. November 1917
 Zeitung
 o.o. - Deutsch - 2 Seiten
Thema: Tageszeitung
 1 Treffer im Text
 ...der Bulgaren nach einen besonderen Pro
 gram sein fleis an die **Front** wert, während...

3
Der Tiroler 18. September 1917
 Zeitung
 Bozeno (Bozen) - Deutsch - 4 Seiten
Thema: Tageszeitung
 3 Treffer im Text
 ...W., der auf Rußland einen Druck ausüben will.

Ao realizar uma busca pela palavra-chave *Brasilien*, por exemplo, o usuário encontra 54.994 resultados. Desses, um contém textos em português, o que pode ser observado na relação de filtros. Trata-se de uma edição do jornal *O Spectador brasileiro*, publicada no Rio de Janeiro em 23 de maio de 1826 (Figura 21).

Figura 21 – Página em português, de *O Spectador brasileiro*, recuperada no ANNO



Ao clicar sobre o jornal escolhido, o usuário é direcionado a todas as páginas daquela edição, estando marcada em vermelho a que contém o resultado da busca. O consulente deve então clicar na página desejada para maximizá-la. A partir da próxima tela, poderá salvar um PDF, copiar o texto extraído pelo OCR, baixar uma cópia em tamanho original e resolução compatível, imprimir a página desejada ou compartilhar via *Facebook*, *Twitter*, *Google+* e e-mail. Outra facilidade permitida pela hemeroteca é copiar o link do resultado da busca, facilitando voltar aos resultados em outra ocasião, sem precisar realizar novamente a busca, como no caso da consulta sobre o Brasil: <http://goo.gl/lnRKBa>.

A página de busca é composta por um mapa interativo (Figura 22), no qual é possível visualizar de onde são os jornais disponíveis no ANNO. Do Brasil, há quatro publicações, com 1925 páginas digitalizadas compreendidas entre os anos 1819 e 1830. São elas *Jornal do Commercio* (684 páginas); *Astrea* (502); *Gazeta do Rio de Janeiro* (457); e *O Spectador Brasileiro* (282).

Figura 22 – Mapa interativo do ANNO com os locais em que foram encontrados os jornais



Fonte: Biblioteca Nacional da Áustria. Disponível em: <<http://anno.onb.ac.at/>>. Acesso em: 15 ago. 2015.

O ANNO oferece uma página de ajuda com dicas sobre como aprimorar a busca: <http://anno.onb.ac.at/suchhilfe.htm> (em alemão).

6.3 AUSTRÁLIA: TROVE

Biblioteca Nacional da Austrália: <http://www.nla.gov.au>

Trove: <http://trove.nla.gov.au/>

A hemeroteca da Biblioteca Nacional da Austrália está localizada dentro do Trove, coleção digital composta por 90 milhões de itens distribuídos em mil bibliotecas, museus, arquivos, galerias e outras organizações do país. Como definido no site de apresentação, Trove é uma comunidade, um conjunto de serviços, um agregador de metadados e um repositório crescente de recursos digitais de texto completo.

Considera-se comunidade, pois conta com a ajuda dos usuários para etiquetar os conteúdos, sugerir correções aos textos extraídos de jornais antigos digitalizados e comentar sobre os itens disponíveis, além de abrigar conteúdos de diferentes organizações. Há ainda um fórum onde podem ser debatidos assuntos referentes ao acervo.

O Trove define-se com um agregador de metadados, mas também armazena conteúdos completos, como, por exemplo, jornais digitalizados. Um dos objetivos da plataforma é oferecer cada vez mais materiais em formato integral.

A digitalização foi iniciada em 2007. A versão beta do sistema foi colocada no ar em 2008 contendo textos digitalizados da coleção de jornais australianos da biblioteca. A seção de jornais digitalizados foi inaugurada oficialmente em 2009 e, no ano seguinte, inserida no Trove.

O serviço de escaneamento é prestado por uma empresa externa, assim como a catalogação dos materiais – não há catalogação por assunto ou indexação. Os equipamentos utilizados para a digitalização pertencem à firma contratada. Após serem escaneadas, as imagens são inseridas no *Newspaper Content Management system*, um sistema localizado na Biblioteca Nacional da Austrália.

Desde o início do processo foram realizadas alterações, como a inserção da digitalização colorida, ao invés da escala de cinzas inicialmente adotada. Atualmente, de acordo com os responsáveis pelo Trove, foram iniciadas as digitalizações realizadas diretamente do jornal impresso, os quais não foram microfilmados.

A decisão da Biblioteca Nacional da Austrália de se tornar líder no acesso à informação produzida sobre e no país chegou a ser questionada, como afirma Holley: “Algumas pessoas podem se perguntar: mas não é isso que o *Google* faz, reunir a informação em um mesmo local? Ainda precisaremos de bibliotecas?” (2010, p. 2). A autora responde apontando as grandes diferenças entre uma instituição e outra.

Bibliotecas são diferentes do *Google* por estas razões: elas se comprometem a prover preservação, curadoria e acesso a seu acervo em longo prazo; elas não têm motivos comerciais para fornecer a informação (considerando várias ações da biblioteca); elas procuram oferecer o acesso universal a todos; e elas são gratuitas para todos. Resumindo: Bibliotecas são permanentes e duram para sempre (HOLLEY, 2010, p. 2, tradução nossa).

Dentro da Biblioteca Nacional da Austrália, há diversas equipes envolvidas com a digitalização do acervo: *Australian Newspaper Digitisation Program* (ANDP) – seis funcionários, sendo quatro bibliotecários; *Australian Newspaper Plan* – dois funcionários, sendo um bibliotecário; *Quality Assurance* – entre seis e 11 funcionários, dependendo do volume e tipo de trabalho; Catalogação – um terço da equipe é contratada; Trove – dois funcionários, ambos bibliotecários; Tecnologia da Informação (TI) – dois funcionários.

O acesso à coleção de jornais digitalizados ocorre no item *Digitised newspapers and more* do Trove, o qual contém 18.138.898 páginas com 178.783.028 artigos disponíveis para busca. O primeiro jornal australiano, *The Sydney Gazette and New South Wales Advertiser*, publicado em Sydney em 1803, pode ser recuperado no Trove.

O Trove é acessado por públicos de diferentes perfis. “Por todos”, como costumam afirmar os responsáveis pelo sistema. “Os usuários das páginas de jornais digitalizados do Trove são desde crianças da escola primária até aposentados. No entanto, a maioria dos usuários procura por informações sobre a história de suas famílias”, descreve a direção do Trove por e-mail.

O número de visitas em julho de 2015 foi de 1.563.952 pessoas de diferentes partes do mundo – sendo 65,6% australianos. As buscas no sistema são superiores a 2,5 milhões por mês.

Não é necessário fazer um cadastro para acessar o acervo, apenas se o usuário quiser fazer uma pasta com material de seu interesse ou contribuir para a melhoria do sistema, sugerindo correções ao texto dos jornais escaneado por meio da tecnologia de OCR ou adicionando etiquetas (*tags*) aos artigos de periódicos. Essas duas alternativas (correção e etiquetagem) são dois diferenciais entre as hemerotecas pesquisadas. As *tags* inseridas pelos usuários são recuperáveis pelo sistema de busca.

O buscador recupera essencialmente conteúdo nacional – há materiais produzidos fora da Austrália, mas que abordam informações sobre o país. Após a instalação do Trove, foram integrados ao sistema serviços de busca que atendiam a outras fontes. Na época do artigo escrito por Holley, em 2010, apenas uma parte do acervo podia ser consultada a partir do OAI-PMH. Apesar disso, muito era recuperado devido a anos de investimento na adoção de padrões mundialmente usados, como MARC, Dublin Core e *Encoded Archival Context*

(EAC). Atualmente, a coleção de jornais antigos é catalogada sob o formato MARC21 seguindo as regras da AACR2 ou do RDA. Também são inseridos no sistema os dados do ISSN de acordo com cada formato de publicação (impresso, microfilmado ou digital), os quais são uma contribuição da base internacional de dados do ISSN.

A maneira como os criadores do Trove encontraram para lidar com os direitos autorais dos jornais foi dando acesso apenas àquelas que já estão em domínio público, ou seja, publicações anteriores a 1954, ou ainda aos periódicos cujos proprietários forneceram autorização. A coleção aberta à consulta na “Web” compreende publicações de 1803 até a metade do século XX.

Dos estimados sete mil títulos existentes na hemeroteca da Biblioteca Nacional da Austrália, apenas mil foram digitalizados até o momento. De acordo com os responsáveis pelo Trove, isso caracteriza o programa de digitalização como um projeto em longo prazo, dependente de orçamento, assim como a participação de organizações e bibliotecas que possam contribuir para aumentar o orçamento disponível.

Ao acessar a página inicial de *Digitised newspapers and more*, uma caixa de busca é o elemento que o usuário percebe com maior destaque, como na Figura 23. Além da busca por palavra-chave, é possível realizar a busca por data, estado ou título na busca simples.

Figura 23 – Página inicial da seção *Digitised newspapers and more* do Trove

The screenshot shows the Trove website interface. At the top, there is a header with the Trove logo and the text "Digitised newspapers and more". Below the header, there are navigation links: "Home", "About", "Browse", and "Help".

The main content area is divided into two sections: "Find an article" and "Find an issue".

Find an article: This section features a search input field and a "Search articles" button. Below the button, there is a link for "Advanced Search".

Find an issue: This section is divided into two sub-sections: "by Title" and "by Date".

by Title: This sub-section lists five titles:

1. The Sydney Morning Herald
2. The Argus
3. Northern Territory Times and Gazette
4. Brisbane Courier
5. The Australian Women's Weekly

A "Show all titles" button is located below the list.

by Date: This sub-section shows a calendar interface for the year 1803. The calendar has columns for months (JAN, FEB, MAR, APR, MAY, JUN, JUL, AUG, SEP, OCT, NOV, DEC) and rows for days of the week (S, M, T, W, T, F, S). The calendar is currently empty, indicating no issues are listed for that year.

by State: This sub-section shows a map of Australia with state abbreviations listed around it: NT, WA, SA, TAS, QLD, NSW, ACT, and VIC.

Fonte: National Library of Australia. Disponível em: <<http://trove.nla.gov.au>>. Acesso em: 15 ago. 2015.

Figura 24 – Página de resultados do Trove

NATIONAL LIBRARY OF AUSTRALIA
Trove

Home About Site news Statistics Tags Help Forum Contact us
It's free and it only takes a minute **Sign up** **Login**

All Books Pictures, photos, objects Journals, articles and data sets Digitised newspapers and more Music, sound and video Maps Diaries, letters, archives Archived websites (1996 - now) People and organisations Lists

Brazil
 Available online Australian content In my libraries **Advanced Search**

Books
view all 111,316 results

- Brazil / Regis St. Louis ... [et al.]
St. Louis, Regis
[Book, Map : 1989-2013]
Possibly online
At 102 libraries
- Brazil / Nick Selby ... [et al.]
Selby, Nick
[Book : 1998]
At 41 libraries
- Brazil / Marion Morrison
Morrison, Marion
[Book : 1988-2012]
View online
At 86 libraries
- Brazil / Steve Goldsworthy
Goldsworthy, Steve
[Book, Map : 2014]
At 31 libraries

Digitised newspapers and more
Showing: 1 - 20 of 311,303 Refine search Sort by: Relevance

BRAZIL.
The Brisbane Courier (Qld. : 1864 - 1933) Thursday 28 March 1867 p 3 Article
...BBAZIL. The mail steamer from Bio brings news that the Emperor of **Brazil** has resolved to increase ...
26 words

BRAZIL.
The Sydney Gazette and New South Wales Advertiser (NSW : 1803 - 1842) Thursday 28 January 1836 p 4 Article
... **BRAZIL** The Frankfort journals contain a long article from the Rio Janeiro, giving the de ... dangerous. It appears that in defiance of all treaties the slave trade is still carried on in **Brazil**, where ... 1121 words
Text last corrected on 4 September 2008 by Itodd

BRAZIL.
The Sydney Morning Herald (NSW : 1842 - 1954) Monday 24 August 1868 p 3 Article
... **BRAZIL** New York, June 15th -Rio Janeiro dates ato to the 9th, and Buenos Ayres to April 27th Lope ...
Buenos Ayres that the United States [allies r1] intended to call a General Assomb) m **Brazil**, which wis ...
134 words

BRAZIL.
The Mercury (Hobart, Tas. : 1860 - 1954) Friday 10 November 1893 p 2 Article
... **BRAZIL** in connection with the Braz. ill.-in revolt, a fatal mistake: has bonn made by the pan... 148 words

Refine your results:

▼ State

- New South Wales (95,163)
- Queensland (66,439)
- Victoria (43,305)
- South Australia (40,705)
- Tasmania (27,333)
- Western Australia (22,323)
- ACT (15,586)
- National (796)
- Northern Territory (663)

▼ Title

- The Canberra Times (AC... (15,508)
- The Telegraph (Brisban... (11,836)
- The Sydney Morning Her... (10,703)
- Daily Commercial News ... (10,416)
- The Mercury (Hobart, T... (8,808)
- The Age (Melbourne, Vi...

Fonte: National Library of Australia. Disponível em: <http://trove.nla.gov.au>. Acesso em: 15 ago. 2015.

Os resultados são apresentados em uma lista com *links* (Figura 24), que podem ser apresentados por relevância e data (crescente e decrescente). Para que o desempenho do sistema mantenha-se alto são apresentados apenas 2 mil resultados por vez.

A busca avançada permite utilizar mais de uma palavra-chave (*All of these words*), uma frase (*The phrase*), a consulta por qualquer uma das palavras listadas (*Any of these words*) ou desconsiderando algumas palavras (*Without these words*). A pesquisa mais completa permite aplicar diversos parâmetros, como escopo da busca (em que local as palavras deveriam aparecer), a data de publicação, o estado de publicação, o nome dos jornais, a categoria do artigo, o tamanho do artigo ou artigos ilustrados. Ainda é possível definir previamente como os resultados devem ser ordenados, se por relevância ou data.

Os itens que compõem a lista são apresentados com informações como palavra-chave, nome do periódico, período, data exata da publicação recuperada, página, categoria, quando o texto foi corrigido pela última vez e por quem. Em seguida, aparece a parte do texto em que o termo pesquisado é citado. Por fim, o número de palavras do texto.

A figura 25 reúne dois exemplos de resultados, um mostrando as etiquetas e correções realizadas por usuários do Trove:

Figura 25 – Detalhe da apresentação dos resultados no Trove

The image shows two search results from Trove, each enclosed in a dashed border. The first result is for an advertisement in 'The Sydney Morning Herald' from Tuesday, 26 February 1850, page 4. The text snippet is: "... WINE. BX **BRAZIL** PACKET, FROM OPORTO DIRECT. MORT AND BROWN HAVE received instructions from Messrs ... 5975 words". Below the snippet, it says "Tagged as: Brazil Packet" and "Text last corrected on 8 May 2013 by anonymous". The second result is for an advertisement in 'The Cornwall Chronicle' from Saturday, 20 February 1836, page 4. The text snippet is: "... ZEALAND. The fine fast-sailing brig **Brazil** Packet, Capt. Crow, will leave for the above port ... 3301 words". Below the snippet, it says "Tagged as: Brazil Packet" and "Text last corrected on 16 April 2012 by acom".

Fonte: National Library of Australia. Disponível em: <<http://trove.nla.gov.au>>. Acesso em: 15 ago. 2015.

Ao escolher uma das opções, o usuário tem acesso à página do jornal. O texto pode ser lido em um PDF na versão original ou em uma coluna à parte, na qual é apresentado o texto que foi recuperado por OCR. Ao lado do botão para visualizar o PDF, o usuário encontra outros para imprimir, visualizar em jpg, texto, informações sobre como citar e como comprar o artigo (Figura 26).

Figura 26 – Resultado de busca empreendida no Trove

The image shows a screenshot of the Trove website interface. At the top, there is a navigation bar with the Trove logo and the text 'Digitised newspapers and more'. Below this, there are search filters and a search bar containing the word 'Brazil'. The main content area displays a list of search results. The first result is highlighted, showing a snippet of text from a newspaper article. The snippet includes the word 'BRAZIL' in bold, followed by a paragraph: 'The situation in Brazil is becoming more critical every day. Admiral Mello is vigorously conducting the bombardment for the revolutionists, and President Peinoto's forces are getting played out.' To the right of this snippet, there is a larger view of the original newspaper clipping, which contains the same text in a larger font. The newspaper clipping also includes the title 'EMPLOYERS' LIABILITY BILL...' and the word 'INTERCOLONIAL...' at the bottom. The interface includes various navigation and search options, such as 'Print', 'PDF', 'JPG', 'TXT', 'Cite', and 'Buy' buttons. There are also links for 'Lists', 'Tagged', 'Comments', and 'Electronically Translated Text'.

Fonte: National Library of Australia. Disponível em: <<http://trove.nla.gov.au>>. Acesso em: 15 ago. 2015.

Entre 2010 e 2011, o Programa de Digitalização de Jornais Australianos mantido pela Biblioteca Nacional da Austrália colocou no ar o chamado “Modelo de Contribuição” que encoraja e facilita a digitalização dos periódicos. O programa está disponível para bibliotecas, grupos comunitários e outras organizações para selecionar e apadrinhar jornais relevantes para toda a sua comunidade e entregá-los ao Trove. A BN gerencia os projetos de digitalização tendo o cuidado de manter um formato padrão visando a preservação em longo prazo, acesso contínuo e “buscabilidade” de textos completos nos artigos de jornais.

O Trove oferece aos usuários uma página de ajuda:

<http://help.nla.gov.au/trove/using-trove/digitised-newspapers/searching-in-newspapers>.

De acordo com a direção do programa, a oferta da página de ajuda diminuiu o volume de mensagens de usuários pedindo suporte e apoio em suas buscas. As dicas começam com o básico, ensinando o usuário a colocar inicialmente uma palavra-chave mais abrangente, “não seja muito específico tão cedo!”, alertam logo de início.

A indicação é refinar a busca após receber os primeiros resultados, usando as facetas oferecidas no menu à esquerda. Facetas é o termo usado para identificar os filtros. São apresentadas as facetas encontradas nos resultados, que podem ser estado, categoria (anúncio, artigo, notícias de família, literatura etc.), título da publicação, década, ilustrações ou número de palavras (menos de 100, entre 100 e 1.000 e com mais de 1.000 palavras).

Todas as mudanças ocorridas no Trove são documentadas no site: <http://trove.nla.gov.au/general/completedlist>.

6.4 BÉLGICA: BELGICAPRESS

Biblioteca Real da Bélgica: <http://www.kbr.be/>

BelgicaPress: <http://opac.kbr.be/belgicapress.php?lang=FR>

O programa de digitalização dos jornais belgas foi iniciado em 2004, quando se criou um plano de digitalização das coleções pertencentes a instituições científicas federais. Para colocar o projeto em prática, a Biblioteca Real da Bélgica (*Koninklijke Bibliotheek van België* - KBR) uniu-se ao Centro de Estudos Históricos sobre Guerra e Sociedade Contemporânea (CEGES). O jornal *Le Soir* (1887-1970) foi o projeto-piloto de digitalização em 2004.

Atualmente, já são mais de 3,75 milhões de páginas digitalizadas de 130 diferentes títulos editados entre os séculos XIX e XX. Porém, em respeito aos direitos autorais, apenas 1,2 milhão de páginas estão *on-line*. As demais podem ser consultadas na sede da KBR, no Centro de Bruxelas.

A equipe da KBR descreve que diferentes projetos de digitalização foram desenvolvidos nos últimos 10 anos abrangendo:

- cerca de 80 diferentes títulos (*DIGIT 01*);
- o jornal *Le Soir*;
- cerca de 10 jornais locais (por exemplo, *Sint-Truiden* e *Dixmuide*);
- cerca de 30 jornais literários;
- quatro títulos digitalizados a partir de microfilmes;
- três jornais que em breve serão inseridos na base de dados (*DIGIT 03*).

A coleção de jornais está dividida em três grandes categorias:

- 1) *Presse quotidienne belge* (1831-1970) – BelgicaPress – seção de jornais digitalizados, incluindo os números “censurados” das duas guerras mundiais.
- 2) *Presse artistique et littéraire* (1871-1958) – digitalização das revistas literárias belgas realizada em 2008 pela Universidade Livre de Bruxelas.

3) *Le projet Gazette de Leyde (1750-1793)* – Projeto realizado em 2007 pela KBR e pelo centro de estudos do 18^o s. de Lyon.

A coleção é formada também por jornais preciosos e conservados na KBR: *Deutsche Brüsseler Zeitung*, *Vlaemisch België* e *Almanach d’Uylenspiegel*.

Os mesmos padrões para digitalização são utilizados desde o início do projeto em 2004, como listado abaixo por Marc D’Hoore, diretor do Departamento de Conservação dos Jornais e Mídias Contemporâneas, e pelas bibliotecárias da KBR Veronique Wese e Maud Henry:

- TIFF, por página, escala de cinza, 300 dpi, *optimized* ("*Original Optimized Scans*");
- PDF, por página contendo uma imagem fraca, 300 dpi, com texto completo oculto;
- PDF, por edição, imagens fracas, 300 dpi, com texto completo oculto;
- JPEG, por página;
- arquivo METS, por edição, estruturado usando esquemas MODS (*Metadata Object Description Schema*) e MIX (*Metadata for Images in XML Standard*);
- arquivo ALTO, por página, contendo o leiaute e texto lido por OCR;
- arquivo Batch.XML, descrevendo os arquivos na coleção digital.

A equipe da KBR estima que sejam digitalizadas ainda 10 milhões de páginas. Isto significa muito trabalho pelos próximos 20 anos. No acervo total há mais de 120 milhões de páginas analógicas a serem escaneadas e gravadas no sistema digital da KBR.

Como tanto a digitalização quanto o processo de OCR são realizados externamente, a equipe na KBR é composta por nove funcionários, sendo cinco no Departamento de Jornais, dois no setor chamado DIGIT, um na parte de tecnologia da informação e um na comunicação. Desses, três são bibliotecários.

A KBR contrata serviços de transporte, digitalização (TIFF, PDF e JPEG), OCR (Abby), criação de metadados METS-ALTO em lotes entregues em HD externos e fitas. Na própria biblioteca são selecionados os títulos para digitalização, realizam-se checagem das coleções, preparação dos lotes, controle de qualidade, gestão do projeto, armazenamento, planejamento para preservação em longo prazo, desenvolvimento da interface do BelgicaPress, *links* para os dados catalogados, redação de notas históricas e manutenção do sistema como um todo.

Desde 24 de abril de 2015, a hemeroteca da KBR conta com um novo mecanismo de busca, apresentado logo na página inicial da seção (Figura 27) no qual o usuário pode buscar artigos por data em um calendário ou utilizando busca por palavra-chave, título, período e idioma – francês, holandês ou inglês. A recuperação se baseia em textos completos, a partir do

escaneamento da página por OCR.

A biblioteca oferece uma página de ajuda, na qual especifica o passo a passo da busca: http://opac.kbr.be/help/jb_FR.PDF.

Figura 27 – Página inicial da coleção de jornais da KBR

The screenshot shows the KBR website interface. At the top, there is a navigation bar with links for 'NL', 'FR', 'EN', 'Contact', and 'Help'. Below this, the KBR logo and 'Royal Library of Belgium' are displayed. A main navigation menu includes 'Catalogue of the KBR', 'Other catalogues', 'Legal deposit', and 'BelgicaPress'. The central area is a search interface with the following elements:

- Search** section:

 - text:** A dropdown menu set to 'All of these words'.
 - newspaper title:** A text input field with a 'Select...' dropdown arrow.
 - period:** Two date input fields (format: dd-mm-yyyy) with a '+' sign between them.
 - language:** A dropdown menu set to 'All'.
 - Search** and **Reset** buttons.

Below the search section is a 'Browse the newspapers by date' section:

- A 'Year...' dropdown menu.
- A 'Newspaper...' dropdown menu.
- A **Reset** button.
- A grid of months:

January	February	March	April
May	June	July	August
September	October	November	December

At the bottom right, there is a copyright notice: '© 2015 Royal Library of Belgium - Legal notice'.

Fonte: Biblioteca Real da Bélgica. Disponível em: < <http://goo.gl/pziJqw>>. Acesso em: 15 ago. 2015.

Os resultados são apresentados em lista de *links* (Figura 28). Ao clicar no item escolhido, o usuário é remetido à imagem da página original do jornal (Figura 29), que pode ser ampliada e/ou salva em PDF. O link da publicação está à disposição. Ao lado esquerdo são listados outros resultados recuperados. À direita, é possível escolher entre resultados apresentados *on-line* ou no acervo físico da KBR. Há um texto disponível sobre o jornal.

Figura 28 – Página com resultados da KRB em forma de lista

The screenshot shows the KBR website interface. At the top, there is a navigation bar with the KBR logo and the text 'Royal Library of Belgium'. Below this, there are links for 'Catalogue of the KBR', 'Other catalogues', 'Legal deposit', and 'BelgicaPress'. On the right side of the header, there are language options 'NL', 'FR', and 'EN', and a 'Contact Help' link. The main content area is titled 'Results for (Brazil)' and shows '18667 pages in 28 newspapers'. Below this, there are two tabs: 'Refine search' and 'New search'. The results are listed in a table-like format with the following columns: newspaper name, page count, and availability status. The results are as follows:

Newspaper Name	Page Count	Availability
1 Handelsblad (Het)	7380 pages	3911 pages available online 3469 pages only available at the Library
2 Meuse (La)	3485 pages	850 pages available online 2635 pages only available at the Library
3 Dernière Heure (La)	2083 pages	7 pages available online 2076 pages only available at the Library
4 Indépendance Belge (L')	1944 pages	562 pages available online 1382 pages only available at the Library

Fonte: Biblioteca Real da Bélgica. Disponível em: < <http://goo.gl/pziJqw>>. Acesso em: 15 ago. 2015.

Figura 29 – Página com um dos resultados da KRB selecionados na lista geral

The screenshot shows the KBR website interface with a selected newspaper result. The page header is the same as in Figure 28. The main content area is titled 'Results for (Brazil)' and shows a search bar and a 'Search in this newspaper' button. Below this, there is a list of newspaper results. The selected result is 'Courrier De L'Escaut (Le) 21-11-1891'. The preview of the newspaper's content is displayed, showing a headline 'L'insurrection au Brésil' and a sub-headline 'LONDRES, vendredi 20. — On mande de Buenos-Ayres à l'agence Reuter: Les dépêches du Brésil annoncent toute la province de Rio-Grande en révolte. 30.000 hommes ont pris les armes contre le dictateur. Le Dr Brazil, chef du comité des républicains, a envoyé un télégramme au baron de Lucena, chef de cabinet du maréchal Fonseca, demandant la démission du républicain.' The preview also shows a section titled 'La Chambre' and a section titled 'FRANCE'.

Fonte: Biblioteca Real da Bélgica. Disponível em: < <http://goo.gl/pziJqw>>. Acesso em: 15 ago. 2015.

Não há indexação ou qualquer tipo de classificação, mas os documentos passam por processo de catalogação, sendo criados campos no MARC21 para título (nome do veículo), data, região, idioma e notas históricas. As buscas são baseadas na aplicação de OCR nos documentos. A equipe estima que haja 99,90% de precisão na recuperação da informação.

Desde que o sistema de busca passou por uma atualização, em maio de 2015, o BelgicaPress foi acessado por 25.454 visitantes, contabilizando mais de 63 mil buscas. Em média, cada visitante navega por 10,5 páginas. Pelo levantamento realizado pela equipe da KBR, os consulentes podem ser divididos em três categorias: estudantes (do ensino fundamental a universitários), pesquisadores científicos e público em geral.

As áreas mais consultadas são ciências sociais (história, filologia, musicologia, ciências políticas, economia, sociologia, jornalismo e comunicação, ciências educacionais), artes cênicas (teatro), fotografia, desenho/charges, exposições, crítica jornalística, reportagens e documentários de TV. Nos primeiros três meses de uso, constatou-se que os visitantes são oriundos de 99 países. Os mais frequentes são belgas (65%). Entre os estrangeiros, 22% vivem nos Estados Unidos, Holanda, França e Alemanha.

6.5 COSTA RICA

Sistema Nacional de Bibliotecas:

http://www.sinabi.go.cr/bibliotecas/biblioteca_nacional.aspx#.UzDfWXfldUg8

Biblioteca Digital / Periódicos: <http://www.sinabi.go.cr/contextual/busqueda.aspx?tipobuscador=4>

Em 2009 foi lançado o portal do Sistema Nacional de Bibliotecas da Costa Rica (SINABI). Considerado parte importante do trabalho da Biblioteca Nacional da Costa Rica, como ressalta Laura Rodríguez Amador, diretora da BN e coordenadora do SINABI, o portal foi iniciado justamente com um jornal e algumas poucas revistas. Ao longo dos últimos anos, outras publicações vêm sendo inseridas na base. Afinal, esta é a única coleção patrimonial de jornais e revistas da Costa Rica.

A digitalização dos jornais é realizada por uma empresa contratada, mas dentro da Biblioteca Nacional, pois determinadas publicações não podem sair da instituição devido ao seu valor patrimonial. Um bibliotecário revisa os arquivos concluídos. Este mesmo funcionário faz a digitalização das revistas. A seleção e a indicação dos jornais a serem digitalizados estão a cargo de outro bibliotecário. Bibliotecários de outros setores colaboram, de acordo com a necessidade e disponibilidade, na parte de controle de qualidade.

O setor de publicações periódicas conta ainda com o trabalho de estudantes universitários, muitos de Biblioteconomia, na inserção de metadados no sistema. Nem todos

os jornais inseridos no site estão indexados no catálogo. Os metadados usados são dia, mês, ano, localidade, temática, título, autoridade e idioma.

No caso das revistas, há ainda informações sobre números anteriores e volume. No portal do SINABI utiliza-se o Dublin Core por meio de um CMS (*Content Management System*, sistema de gestão de conteúdos). Já no catálogo da BN, os metadados adotados são os do MARC21. A catalogação dos jornais é realizada pela informação geral do título, não há catalogação por edição.

O material é digitalizado colorido com 300 dpi de resolução por um *scanner* aéreo com luz fria para não deteriorar as páginas. Porém, a versão inserida no portal SINABI tem 150 dpi e é apresentada em preto e branco no formato PDF. Esse padrão é seguido desde o início do projeto. Há alteração no modelo somente quando o jornal digitalizado está em péssimo estado de conservação. Procura-se uma alternativa para manter a qualidade da cópia.

Até setembro de 2015 já foram digitalizados 518 títulos de jornais, englobando 115.000 edições, além de 40 títulos de revistas, o equivalente a 1.848 números. Como avalia Rodríguez Amador, ainda há muito trabalho a ser realizado. No caso dos jornais, para alguns ainda faltam as autorizações dos proprietários, especialmente os posteriores a 1950. Já com as revistas, as limitações são econômicas e de pessoal. “Vamos por etapas, fazendo um pouco a cada ano”, afirma com otimismo a diretora do SINABI, apesar dos cortes que o projeto sofreu no ano de 2015.

O trabalho de digitalização foi iniciado com jornais do século XIX, mas atualmente grande parte da coleção da primeira metade do século XX já foi coberta pelo projeto. “De jornais muito importantes da Costa Rica, e dos quais já recebemos autorização, conseguimos cobrir de 1970 a 1990.” No caso das revistas, o processo é um pouco mais lento, mas boa parte do acervo do século XIX e algumas edições do século XX já estão no sistema.

A busca no portal do SINABI pode ser realizada por título, ano, dia, mês e assunto (Figura 30) – ou uma combinação de algumas dessas categorias. É possível inserir palavras-chave, mas haverá mais chance de recuperação da informação ao se usar as categorias oferecidas na página de busca mencionadas anteriormente. Antes de realizar a busca, o usuário já pode escolher como pretende visualizar os resultados (se ordenados por relevância, título, tipo, autor, data de publicação, tamanho ou localização).

Não há uso de OCR. Assim, a busca por palavra-chave se limita aos títulos dos jornais. O sistema sugere outras palavras-chave quando não recupera nada. Os resultados são apresentados em forma de lista no formato PDF (Figura 31).

Figura 30 – Página de busca nos periódicos costa-riquenos

Inicio | Quejas y sugerencias | Mapa del sitio | Acerca de SINABI | Ingresar

SISTEMA NACIONAL DE BIBLIOTECAS COSTA RICA

Pregúntele a la Biblioteca | CHAT Inactivo

en: Todos | Buscar

Biblioteca digital | Exhibiciones y subportales | Niños y jóvenes | Actividades | Servicios | Bibliotecas | Catálogo | Contáctenos

[No se encontraron resultados con su criterio de búsqueda]

Buscador de periódicos

Inicio

Buscador

Buscar: Ordenamiento: Buscar

Día de publicación: Mes de publicación: Año de publicación o grabación:

Tema: Administración (6) Arquitectura (138) Artes plásticas (1321) Bibliotecología (146)

Provincia: Alajuela (418) Cartago (1515) Guanacaste (4640) Heredia (541)

Quiso decir: [Brasil](#)

Lo más reciente en el portal

Libros 09-07-2015
Publicado: 1963 Autor: Costa Rica
Ministerio de Transportes, Instituto Geográfico de Costa Rica
Reproducciones científicas, una expedición y legislación de la Isla del Coco. Parte 2

Libros 09-07-2015
Publicado: 1963 Autor: Costa Rica
Ministerio de Transportes, Instituto Geográfico de Costa Rica

Usuarios

- Ciudadanos
- Investigadores

Suscripciones

- Actividades
- Boletines
- Novedades en Biblioteca Digital

Fonte: SINABI. Disponível em: <<http://goo.gl/vLpz5S>>. Acesso em 15 jan. 2016.

Figura 31 – Lista de resultados obtidos no SINABI

Resultados 1 - 10 de un total de 10000 Duración 00:00:03.8576881 segundos

Quiso decir: [Brasil](#)

Documento	Tamaño
[PDF] De la imprenta al periódico: historia de la comunicación social en Costa Rica 1821-1850 - Publicado 10/02/2009 Historia del periodismo en Costa Rica desde la independencia de Costa Rica hasta 1850 /biblioteca digital/LIBROS COMPLETOS/Vega Jimenez Patricia/De la imprenta al periodico Completo.pdf idioma -	- 4,66 MB
[PDF] Diario de Costa Rica - Publicado 24/02/2011 ku-Diario de Costa Rica_23 nov_1940 /biblioteca digital/periodicos/diario de costa rica/diario de costa rica 1940/ku-Diario de Costa Rica_23 nov_1940.pdf idioma -	- 2,81 MB
Diario de Costa Rica - Publicado 24/01/2011 Inició el 1 de enero de 1885, con el respaldo de quienes impulsaron el Mentor costarricense, como el Dr. Castro Madriz y Joaquín Bernardo Calvo Mora, quien fungió como su administrador. Entre los colaboradores estaban Juan Fernández Ferraz, Angel Anselmo Castro y Rafael Montúfar. Incluía información sobre diversos tópicos, así como anuncios. Según Adolfo Blen, con este periódico "quedo forjado en una manera definitiva el diarismo en Costa Rica ya que era una verdadera necesidad" /biblioteca digital/periodicos/Diario de Costa Rica1.aspx idioma -	- 14,55 KB
Costa Rica - Publicado 23/07/2010 Periódico bisemanal editado por Rafael Carranza. Su primer número se publicó el 17 de junio de 1893. De carácter noticioso, literario, y según el mismo editor con un interés nacionalista /biblioteca digital/periodicos/Costa Rica.aspx idioma -	- 19,29 KB

Fonte: SINABI. Disponível em: <<http://goo.gl/vLpz5S>>. Acesso em 15 jan. 2016.

Mensalmente registram-se 5 mil acessos ao portal, sendo realizadas mil buscas no sistema. Rodríguez Amador descreve esses visitantes como pesquisadores, estudantes universitários e de ensino médio, mas também uma grande parcela de público em geral. Para a diretora, o importante no trabalho desempenhado pelo portal do SINABI é agilizar o acesso a informações que por muitos anos estavam restritas a um pequeno número de usuários, assim como proteger coleções preciosas que antes podiam se deteriorar pelo uso massivo.

6.6 ESPANHA: HEMEROTECA DIGITAL

Biblioteca Nacional de España - <http://www.bne.es/>

Hemeroteca Digital - <http://hemerotecadigital.bne.es/>

A Hemeroteca Digital espanhola faz parte do projeto *Biblioteca Digital Hispánica*, que tem como objetivo a consulta e a difusão pública do patrimônio bibliográfico espanhol conservado na Biblioteca Nacional.

O site da Hemeroteca Digital da Biblioteca Nacional da Espanha (BNE) foi inaugurado em março de 2007, mas as primeiras tentativas de digitalização da coleção de jornais iniciaram-se ainda em 1997, quando foram microfilmados ou digitalizados a partir dos microfilmes 187 títulos do século XIX e meados do XX.

Somente em 1999 o trabalho foi organizado de forma sistemática, motivado pela necessidade de preservar o acervo, assim como de facilitar a consulta nessas publicações especialmente frágeis que são jornais e revistas antigos.

Em 2006, a digitalização tomou dois rumos distintos. Se por um lado se continuou a reprodução sistemática da imprensa contemporânea – esta disponível somente na sede da Biblioteca Nacional espanhola –, por outro se começou a integrar jornais e revistas antigos livres de direitos autorais à coleção. Tais publicações puderam então ser consultadas na rede. A seleção dos jornais e revistas contemplou títulos de temática variada, representando a diversidade da coleção, que engloba imprensa satírica, humorística, política, técnica, profissional, religiosa, ilustrada, esportiva, literária e educativa produzida entre o século XVIII e XX.

De acordo com Rodríguez Fuentes (2011, p. 23), coordenadora da Hemeroteca Digital e integrante da diretoria do Departamento de Revistas da BNE, a seleção baseou-se principalmente em três critérios: conservação, uso e tipo de material.

Inicialmente houve no grupo uma discussão sobre a conveniência ou não da digitalização como ação de conservação frente à ideia já estabelecida de que a microfilmagem

é o sistema que melhor assegura a sobrevivência de conteúdos originais.

Ao mesmo tempo foram se estabelecendo as ideias de que a digitalização apoia os programas de preservação, já que ao não usar os originais, estes não são submetidos à deterioração decorrente da manipulação. A seleção dos materiais a serem digitalizados contemplou jornais e revistas em estado ruim de conservação, já retirados da consulta, exemplares únicos de grande interesse e publicações submetidas frequentemente ao uso e à cópia. As estatísticas de uso foram muito importantes nessa fase, como observa Rodríguez Fuentes (2011, p. 24).

Jornais e revistas espanhóis antigos foram selecionados por trazerem uma visão completa dos acontecimentos políticos, sociais e culturais da época em que foram produzidos. Foram escolhidas tanto revistas especializadas/temáticas quanto revistas ilustradas de notícias.

Tão logo se definiu o objetivo da coleção que seria colocada no ar, foi realizada uma revisão de todos os materiais pertencentes ao acervo da seção de publicações periódicas da Biblioteca Nacional da Espanha. Rodríguez Fuentes (2011, p. 25) afirma terem sido revisados os catálogos e as bases de dados, além de página a página de todos os exemplares possíveis. Nesse trabalho notou-se quais jornais e revistas tinham problemas de faltas de páginas, manchas, rasgos e mutilações. Essas informações foram coletadas para um possível intercâmbio com acervos de outras instituições. Alguns dos exemplares passaram por restauração.

Inicialmente o resultado da digitalização era obtido em preto e branco, mas atualmente é colorido. Já foram inseridos no acervo digital 1.275 títulos, ou seja, 17.871.777 páginas. Há materiais datados entre 1683 e 2009. Rodríguez Fuentes (2015) afirma ser impossível digitalizar a coleção completa, pois são mais de 160 mil títulos, mas espera-se digitalizar pelo menos mais 500 títulos nos próximos anos. A digitalização é realizada por uma empresa externa. Na biblioteca atuam quatro bibliotecários e três auxiliares, que se ocupam da seleção, revisão e difusão da hemeroteca.

Os profissionais catalogam o acervo utilizando o formato MARC21. Indexação, porém, não é realizada. As buscas baseiam-se na aplicação de OCR.

Em todas as bibliotecas que se estão investindo em projetos de digitalização, analisa-se a conveniência de aplicar as técnicas de OCR a jornais antigos e a exatidão dos resultados obtidos. Não se pode perder de vista que o OCR sempre será uma busca textual e não controlada, o que, ainda que sua transcrição fosse exata, seguiria entregando-nos resultados errados, e muitos “ruídos” em nomes comuns e datas (RODRÍGUEZ FUENTES, 2011, p. 27).

Apesar de os resultados oferecidos pelo uso de OCR estarem ainda longe da perfeição, como afirma Rodríguez Fuentes (2011, p. 27), “esta é a única ferramenta de busca

de conteúdo que podemos oferecer a nossos usuários”. A média de reconhecimento nos testes empreendidos pela BNE tem alcançado 86%.

Nas buscas na hemeroteca da BNE, de acordo Rodríguez Fuentes, além do OCR, conta-se com operadores booleanos, além de busca por âmbito geográfico e cronológico. “Atendemos a todas as consultas e sugestões dos pesquisadores, os ajudamos a refinar suas buscas e em breve será possível realizar uma busca por materiais.” Desde 2012, se cumprem os padrões OAI e EUROPEANA.

Os usuários da Hemeroteca Digital espanhola são majoritariamente universitários e pesquisadores. Em média, 30 mil pessoas utilizam o serviço, registrando 57 mil visitas mensais e aproximadamente 800 mil páginas vistas.

Existem três abas na página principal. Uma com critérios de busca, outra para os resultados e uma terceira para detalhes do que foi recuperado (Figura 32). Na aba em que se realiza a busca, é possível digitar palavras-chave ou selecionar itens da coleção por título, ano, período ou âmbito geográfico. Utiliza-se um dos recursos ou, se o usuário preferir, pode combiná-los.

Figura 32 – Página de busca da Hemeroteca Digital da Biblioteca Nacional da Espanha

BIENVENTIDA NACIONAL DE ESPAÑA
BNE
HEMEROTECA DIGITAL
Biblioteca Digital Hispánica

Consulta Resultados Detalles Listado de títulos | Novedades | Próximamente | Otras hemerotecas digitales | Contacto | Ayuda

Documentos que contengan las palabras en

contengan las palabras en

contengan las palabras en

Título	Ámbito geográfico	Año
<input type="checkbox"/> La Abeja (Barcelona)	<input type="checkbox"/> Álava	<input type="checkbox"/> 1683
<input type="checkbox"/> La Abeja del Turia	<input type="checkbox"/> Albacete	<input type="checkbox"/> 1685
<input type="checkbox"/> Abeja española	<input type="checkbox"/> Alicante	<input type="checkbox"/> 1689
<input type="checkbox"/> La Abeja manresana	<input type="checkbox"/> Almería	<input type="checkbox"/> 1690
<input type="checkbox"/> AC. Documentos de actividad contemporánea	<input type="checkbox"/> Argentina	<input type="checkbox"/> 1691
<input type="checkbox"/> Academia heráldica	<input type="checkbox"/> Asturias	<input type="checkbox"/> 1692
<input type="checkbox"/> La Acción (Madrid, 1916)	<input type="checkbox"/> Ávila	<input type="checkbox"/> 1693
<input type="checkbox"/> Acción española (Madrid)	<input type="checkbox"/> Badajoz	<input type="checkbox"/> 1694
<input type="checkbox"/> La Acción social (Zaragoza)	<input type="checkbox"/> Baleares	<input type="checkbox"/> 1696
<input type="checkbox"/> Acción socialista (Madrid, 1914)	<input type="checkbox"/> Barcelona	<input type="checkbox"/> 1719

Fecha entre el 01 / 01 / 1683 >> y el 31 / 12 / 2015

Buscar Limpiar Recuperar Guardar Cargar

Documentos ordenados:

Mostrar: Títulos Ejemplares Páginas

CrashZandora v4.17

Fonte: Hemeroteca Digital da BNE. Disponível em: <<http://hemerotecadigital.bne.es/>>. Acesso em: 16 jan. 2016.

Um detalhe interessante: o sistema permite que se guarde a busca realizada por meio das teclas *cargar* e *guardar*. O formato dos resultados varia de acordo com o tipo de filtro usado para realizar a busca. Normalmente, são apresentados em forma de lista, na qual constam dos seguintes dados: reprodução da página, nome do periódico, número da edição e número da página (Figura 33). Em alguns deles, aparece a cidade onde o exemplar foi publicado e o ano de veiculação. Há ainda um trecho da frase em que o termo buscado é citado. É possível ordenar por mais novos ou mais recentes, relevância, de A a Z ou de Z a A.

Figura 33 – Página de resultados da Hemeroteca espanhola



The screenshot shows the search results page for the term "Brazil" on the Hemeroteca Digital BNE website. The page header includes the BNE logo and navigation links. The search results are displayed in a list format, with each entry showing a thumbnail of the document page, the title of the newspaper, the date, issue number, and page number. The first three results are from "El Español (Londres)" dated 30/7/1810, n.º 3, pages 124, 162, and 76. Each result includes a "Ver ejemplar" link and social media sharing icons for Facebook and Twitter.

Bienvenido | Benvingut | Ongi etorri | Benvido | Welcome | Bienvenue

Biblioteca Digital Hispánica

Consulta Resultados Detalles Listado de títulos | Novedades | Próximamente | Otras hemerotecas digitales | Contacto | Ayuda

Resultados de su consulta

Ejemplares y páginas que contengan "Brazil" en el texto de las páginas. Ordenados de más antiguo a más reciente.

Página 1 de 154. Resultados: 1.540. Ordenados de más antiguo a más reciente

← anterior | 1 | siguiente →

El Español (Londres). 30/7/1810, n.º 3, página 124.

28b' sion de que su gobierno trataba de • abandonar la madre patria, y separarse de la causa de su cautivo rey. Tal fue la insurrección de México^ que acabó en la deposición y arresto del Virey, a quien sospechaban ; y tal parece haber sido la causa de los últimos alborotos de la Paz. Liniers, no obstante ser el libertador de Buenos Ayres, no ha podido mantener su autoridad en aquel pueblo, contra la sospecha de ser desafecto a España ; y la princesa del Brazil, si es verdad lo que corre, no

Ver ejemplar • Detalles •  

El Español (Londres). 30/7/1810, n.º 3, página 162.

habían opuesto. Según las noticias, la revolución se ha extendido desde Cartagena al Brazil; la Havana

Ver ejemplar • Detalles •  

El Español (Londres). 30/8/1810, n.º 5, página 76.

nombrado» se establecen, n s^ber: Por quintal métrico: Lo» algodones del Brazil. Cayena, Surinam

Ver ejemplar • Detalles •  

Fonte: Hemeroteca Digital da BNE. Disponível em: <<http://hemerotecadigital.bne.es/>>. Acesso em: 16 jan. 2016.

Em cada um dos resultados há dois *links*: um é “ver exemplar” (que direciona para a página digitalizada, de onde se pode baixar um PDF) e o outro é “detalhes” (por meio do qual o usuário tem acesso ao texto para ser copiado), além dos ícones do *Twitter* e do *Facebook*, para quem quiser compartilhar determinado item.

Ao clicar no link da página escolhida, o usuário é direcionado para a página do periódico na qual são mencionadas as palavras “buscada” e “encontrada”. Em um menu à esquerda, o leitor pode acessar outras páginas da mesma edição (Figura 34). O refinamento da busca pode ser feito somente antes de iniciá-la. Uma vez realizada, o usuário vê apenas os resultados obtidos, sem poder filtrá-los novamente a partir deste ponto da consulta.

Figura 34 – Página recuperada na Hemeroteca Digital espanhola



Fonte: Hemeroteca Digital da BNE. Disponível em: <<http://hemerotecadigital.bne.es/>>. Acesso em: 16 jan. 2016.

6.7 ESTADOS UNIDOS: *CHRONICLING AMERICA*

Library of Congress: <http://www.loc.gov/>.

Chronicling America: <http://chroniclingamerica.loc.gov/>

Se o usuário decidir entrar na área de jornais digitalizados da Biblioteca do Congresso (*Library of Congress*, LC), ele terá primeiro que clicar em *News and Journalism* e depois em *Chronicling America: Historic American Newspaper (CA)* – <http://www.loc.gov/>. O caminho não é exatamente intuitivo. Erra-se até se chegar efetivamente à coleção de jornais. Há, no entanto, um link direto para quem já sabe da existência do acervo: <http://chroniclingamerica.loc.gov>.

O CA é produzido pelo *National Digital Newspaper Program (NDNP)*, uma parceria entre *National Endowment for the Humanities (NEH)* e *Library of Congress*. Composto por

jornais antigos e páginas selecionadas, é uma derivação do catálogo criado por instituições estatais durante o programa de jornais dos EUA da NEH entre 1980 e 2007. As primeiras digitalizações foram feitas a partir de microfimes, por questões de custos e eficiência.

A versão beta do CA foi liberada para um grupo limitado em 2006, sendo aberta ao público em março de 2007. A construção do repositório foi baseada em três motivações: apoiar a digitalização de jornais com significado histórico, facilitar o acesso público via “Web” e prover a preservação em longo prazo. Anualmente a NDNP seleciona organizações interessadas em fechar parcerias para escaneamento de coleções.

A coleção digitalizada compreende 1.800 títulos do período entre 1836 e 1922. O objetivo é inserir jornais de cada estado norte-americano. Dos 50 existentes, 40 já foram representados. Oito funcionários da LC estão envolvidos no programa de desenvolvimento do CA, no fluxo de produção, trabalhando com parceiros, realizando controle de qualidade, transferência de dados e desenvolvimento de softwares. Outros funcionários da LC atuam na gestão dos dados, na manutenção do site e no desenvolvimento de ferramentas. Cada instituição parceira envolvida no processo costuma empregar de duas a cinco pessoas.

Cada objeto digitalizado tem tamanho médio de 49 Mb. As páginas são gravadas em documentos únicos – utilizando os seguintes formatos e tecnologias: TIFF, JP2, PDF, OCR ou XML. Todas as orientações sobre o processo técnico de digitalização são apresentadas no site <http://www.loc.gov/ndnp/guidelines/>.

Na página inicial (Figura 35), tem-se acesso a um sistema de busca simples, por meio do qual se pode escolher o estado em que foi publicado o jornal e o período. Há também uma área para busca por palavras-chave.

Na busca avançada, pode-se pesquisar por estado, título do jornal, ano, período, idioma (inglês, francês, espanhol e alemão), palavras-chave, além de outras combinações de palavras (*with any of the words*, *with all of the words* ou *with the phrase*).

A possibilidade de usar esses metadados ajuda os usuários a refinar resultados e a buscar publicações específicas. De acordo com Deborah Thomas, gerente de programas da NDNP/LC, pesquisadores, educadores e estudantes costumam ser o público mais frequente, assim como realizadores de pesquisas genealógicas.

O conteúdo do CA é catalogado por título, data, edição e sequência de página. Cada título está identificado no catálogo da LC. São utilizados os seguintes formatos: MARCXML, METS, MODS, ALTO, PREMIS, MIX.

Os resultados são apresentados em formato JPEG, PDF e em formato de texto gerado a partir de OCR, que pode ser facilmente copiado. Seguem-se os padrões OAI de preservação.

No site, é possível ter acesso a particularidades do sistema em <http://chroniclingamerica.loc.gov/about/api/>, assim como tirar dúvidas na seção de ajuda: <http://chroniclingamerica.loc.gov/help/>. O material recuperado por ser enviado para diferentes redes sociais, como *Facebook*, *Twitter*, *Pinterest* e *Google+*.

Figura 35 – Página inicial do *Chronicling America*

Fonte: Library of Congress. Disponível em: <<http://chroniclingamerica.loc.gov/>>. Acesso em: 17 jan. 2016.

Para ler um dos resultados, basta o usuário clicar sobre a página selecionada (Figura 34), a qual será aberta dentro de um visualizador, onde há botões para aumentar ou diminuir o tamanho da figura, opção de navegar pela edição completa, exibir todas as páginas ou exibir todas as edições disponíveis daquele periódico (Figura 36) – neste caso, o leitor é remetido para outra página, onde há um calendário, no qual ele irá escolher a data desejada.

Há ainda *links* para o texto do documento recuperado, o PDF ou um JPG, com o respectivo tamanho. Além disso, o usuário pode recortar uma parte do conteúdo, que será

apresentada em forma de imagem.

Figura 36 – Página recuperada no *Chronicling America*

Search Pages Advanced Search All Digitized Newspapers 1836-1922 US Newspaper Directory, 1690-Present

All states + from 1836 to 1922 + enter one or more search words 60

Print Subscribe Share/Save Give Feedback

Evening star., December 03, 1922, Image 1
About Evening star. (Washington, D.C.) 1854-1972

Image provided by: Library of Congress, Washington, DC

Image: 1 of 96. Page → All Pages ← Issues → All Issues Text PDF JP2 (4.2 MB)

The Sunday Star.
WASHINGTON, D. C., SUNDAY MORNING, DECEMBER 3, 1922.—NINETY-SIX PAGES.

Sees Labor War on Road To Force U. S. Ownership
Hooper Charges Wide Campaign to Stir Hatred.
Effort Made to Undermine Courts, It Is Charged.

RECEIVES PLEA AT PARLEYS
Delegates From Many States Promise Aid in Campaign.
SEE SIGNS OF SUCCESS FOR PROPOSED MEASURES.
Administration Trying to "Thunder" Leaders in Decision.

CHRISTMAS SHOPPING

AND WE HAD THE FIRST CHANCE OF EVERYTHING

THE ONLY WAY THE COUPERS WERE ALL SO RICE TO

GET NO MORE LIST

By the Associated Press.
CHICAGO, December 2.—Sen. W. Hooper, chairman of the Railroad Labor Board, in a speech before the Illinois Bar Association here tonight, charged that many railway union leaders are waging warfare against the roads with the purpose of destroying the industry and forwarding a political campaign for governmental ownership.

"A large portion of the union managers which seek to obtain better strikes on the railroads, their management and their policies," he said, "These criticisms are not dealing in matters of labor's contemptuous behavior toward the roads and the employees, but they cover every ground of attack that might be made by those outside of railroad employment."

"Nothing is left unaided that seems to be calculated to further govern-

Fonte: Library of Congress. Disponível em: <<http://chroniclingamerica.loc.gov/>>. Acesso em: 17 jan. 2016.

Figura 37 – Opções de visualização do material recuperado no CA

Evening star., December 03, 1922, Image 1
About Evening star. (Washington, D.C.) 1854-1972

Image provided by: Library of Congress, Washington, DC

Image: 1 of 96. Page → All Pages ← Issues → All Issues Text PDF JP2 (4.2 MB)

The Sunday Star.

Fonte: Library of Congress. Disponível em: <<http://chroniclingamerica.loc.gov/>>. Acesso em: 17 jan. 2016.

Ao fazer esta navegação por um item recuperado, o usuário perde o resultado da busca realizada, podendo reencontrá-la apenas com o uso do botão voltar. Uma vez empreendida a busca, não é possível refiná-la.

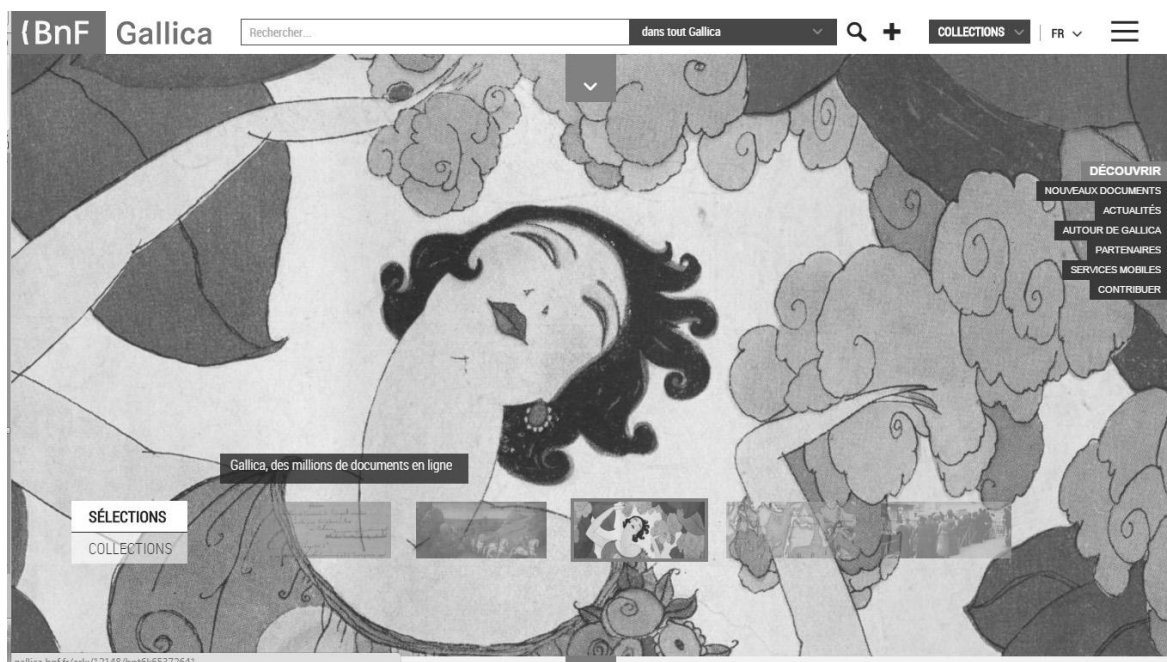
6.8 FRANÇA: GALLICA

Biblioteca Nacional da França: <http://www.bnf.fr>

Gallica / Presse et revues: <http://gallica.bnf.fr/html/presse-et-revues/presse-et-revues>

Assim como a Trapalanda e o Trove, a Gallica reúne não apenas jornais, mas uma ampla coleção de documentos digitalizados, sendo a maior parte pertencente ao acervo da Biblioteca Nacional da França (BNF) – são 3.480.461 documentos, 605.942 livros, 88.614 mapas, 56.877 manuscritos, 1.166.736 imagens, 1.499.434 jornais e revistas, 28.472 partituras, 34.386 gravações de som. A Figura 38 mostra a página inicial da Gallica.

Figura 38 – Página inicial da Gallica, a Biblioteca Digital da BN francesa



Fonte: Biblioteca Nacional da França. Disponível em: <<http://www.bnf.fr>>. Acesso em: 12 jan. 2016.

A Gallica está no ar desde o fim de 1997. A cada semana recebe milhares de novas informações relacionadas à língua francesa e que tratam sobre diversas áreas do conhecimento. As digitalizações são realizadas tanto pela BNF quanto pelos próprios parceiros. Dos 70 mil documentos digitalizados anualmente, um terço é realizado por bibliotecas parceiras. Mesmo os documentos digitalizados fora da BNF recebem o tratamento bibliográfico adotado pela Biblioteca Nacional e os metadados são inseridos no catálogo geral, o que facilita a busca.

Inicialmente as páginas eram digitalizadas em preto e branco com 300 dpi de resolução. Atualmente, usa-se padrão 400 dpi colorido, segundo Sanz (2012, p. 4).

Antes de serem digitalizados, os materiais passam por técnicas de preparação – o que não quer dizer restauração, mas limpeza, alisamento das páginas, soltura das páginas quando encadernadas.

A seção de jornais e revistas chama-se *Presse et revues* (Figura 39) e foi inicialmente colocada no ar em 2006. Estão disponíveis para consulta *on-line* as publicações livres de direitos autorais, ou seja, veiculadas há mais de 70 anos. A primeira coleção digitalizada foi a dos jornais incluídos na chamada Terceira República, de 1870 a 1940, considera por Vallas (2012, p. 5), do Departamento de Conservação da BNF, como “a idade de ouro da imprensa, mas a idade negra do papel”.

Figura 39 – Página inicial da seção *Presse et revues* da Gallica

The screenshot shows the Gallica website interface. At the top left, there is a search bar with the text "Rechercher..." and a dropdown menu for "FR". Below the search bar, there are navigation options: "dans tout Gallica", "COLLECTIONS", and a plus sign icon. The main content area is titled "Presse et revues" and contains a large text block with the following text:

Si l'on considère *La Gazette de Théophraste Renaudot* fondée en 1631 comme le premier journal français, c'est avec *Le Journal de Paris* que s'ouvre en 1777 l'ère de la presse d'information.

1789 instaure la liberté de la presse. Les feuilles éclosent en nombre. (*Le Père Duchesne*, *L'Ami du Peuple*...) avant d'être à nouveau réprimées. Le combat contre la censure agite tout le XIXe siècle qui s'achève par un « âge d'or de la presse » à la faveur de la Loi sur la liberté de la presse de 1881.

Des 1836, les formules novatrices du *Sigée* et de *La Presse* ouvrent la marche de cet essor. Les titres généralistes retranscrivent les partitions politiques et alimentent le débat. La presse satirique (*Le Charivari*, *Le Journal pour rire*...) se développe également qui à travers caricatures et portraits-charges dénonce le pouvoir et s'attaque par l'humour un public fidèle. Plus étroites que les journaux et d'une périodicité moindre, des publications prestigieuses (*La Revue des deux mondes*, *La Revue de Paris*) ou émanant de sociétés savantes parfois locales reflètent quant à elles l'intérêt croissant de l'époque pour

Below the text, there are several thumbnail images of newspaper front pages, each with a corresponding label:

- L'Ouest-Eclair**: Presse locale et régionale
- L'Aurore**: Les principaux quotidiens
- Bellica**: Presse par thématiques
- L'Hindou**: Presse locale et régionale

At the bottom left, there is a sidebar with the following menu items:

- Les principaux quotidiens
- Presse par thématiques
- Presse locale et régionale
- Presse par pays ou zones géographiques
- Journaux de tranchées
- Journaux clandestins de la Résistance
- Revue savantes

A coleção de jornais abrigada na Gallica está dividida em cinco grandes categorias: Principais jornais diários; Imprensa geral e especializada; Imprensa por países ou zonas geográficas; Revistas por região; e “relatos das trincheiras” (*Feuilles de tranchées*). Dentro de cada seção, os jornais são organizados em ordem alfabética.

A Gallica permite tanto a busca simples (Figura 40), por meio da inserção de palavras-chave em uma caixa de texto, quanto a busca avançada. Na busca simples, pode-se selecionar a seção da coleção em que se pretende pesquisar, como a de jornais e revistas (*Presse et revues*). O mecanismo de busca é usado em qualquer seção da biblioteca, seja para buscar documentos presentes somente no acervo físico quanto os documentos digitalizados.

Figura 40 – Busca simples para qualquer seção do Gallica



Fonte: Biblioteca Nacional da França. Disponível em: <<http://www.bnf.fr>>. Acesso em: 12 jan. 2016.

A busca avançada inclui operadores booleanos, além de a opção para se selecionar filtros como ano, título, língua, tipo de documento, assunto, biblioteca onde o objeto está hospedado ou em qual site (Gallica ou site parceiro). Quando a busca é realizada por título, por exemplo, o resultado traz uma tabela com os anos de publicação do jornal escolhido (Figura 41). A partir desta listagem, pode-se ter acesso a um calendário de determinado ano e navegar pelo dia escolhido. O leitor chegará então à página desejada, que poderá ser ampliada, gravada em PDF. O usuário tem a possibilidade de copiar um link com a localização da página.

O site permite ainda realizar uma busca na página, solicitar uma impressão ou apontar alguma incorreção que tenha sido observada. Ao clicar em *Aide* (ajuda), são apresentadas ao usuário diferentes maneiras de visualização, assim como informações adicionais sobre o que está sendo exposto. Há ainda opções de navegação pela edição selecionada.

Figura 41 – Resultado da busca realizada por título na Gallica

Accueil > Années disponibles
Le Figaro (Paris. 1854)

RECHERCHER

Rechercher dans tous les numéros

Informations détaillées

89 années disponibles - 28900 numéros

				1854 41 numéros	1855 52 numéros	1856 103 numéros	1857 107 numéros	1858 106 numéros	1859 105 numéros
1860 105 numéros	1861 104 numéros	1862 104 numéros	1863 105 numéros	1864 104 numéros	1865 105 numéros	1866 135 numéros	1867 371 numéros	1868 364 numéros	1869 378 numéros
1870 362 numéros	1871 298 numéros	1872 364 numéros	1873 369 numéros	1874 348 numéros	1875 364 numéros	1876 387 numéros	1877 382 numéros	1878 376 numéros	1879 368 numéros
1880 368 numéros	1881 368 numéros	1882 368 numéros	1883 366 numéros	1884 366 numéros	1885 365 numéros	1886 365 numéros	1887 365 numéros	1888 366 numéros	1889 366 numéros
1890 366 numéros	1891 367 numéros	1892 372 numéros	1893 369 numéros	1894 371 numéros	1895 370 numéros	1896 368 numéros	1897 366 numéros	1898 366 numéros	1899 395 numéros
1900 366 numéros	1901 365 numéros	1902 367 numéros	1903 369 numéros	1904 372 numéros	1905 368 numéros	1906 375 numéros	1907 367 numéros	1908 368 numéros	1909 370 numéros
1910 389 numéros	1911 369 numéros	1912 366 numéros	1913 367 numéros	1914 364 numéros	1915 361 numéros	1916 366 numéros	1917 365 numéros	1918 365 numéros	1919 343 numéros

Fonte: Biblioteca Nacional da França. Disponível em: <<http://www.bnf.fr>>. Acesso em: 12 jan. 2016.

Em setembro de 2015, a biblioteca digital da Biblioteca Nacional francesa passou por uma remodelação integral, inclusive com alteração do logotipo e relançamento do site com novo visual, menos colorido. Algumas páginas, antes disponíveis em vários idiomas, por ora apresentam-se apenas em francês.

Para aqueles que têm dificuldades de navegar pela biblioteca digital, há uma página de ajuda: <http://gallica.bnf.fr/html/und/rechercher-dans-gallica>. Outra opção é trocar mensagens com a equipe via blog: <http://gallica.bnf.fr/blog>.

Uma novidade é a possibilidade de baixar um aplicativo no tablete ou celular para navegar no acervo da Gallica. O sistema de busca é mais simplificado do que o da versão na “Web”, mas os resultados são os mesmos. Navega-se pelas páginas recuperadas, sendo possível, por meio de *zoom*, ler os materiais apresentados (Figura 42), assim como salvá-los como imagem ou enviá-los por e-mail.

Figura 42 – Visualização de uma página recuperada no app da Gallica



Fonte: Biblioteca Nacional da França (2016).

6.9 LIECHTENSTEIN: ELIECHTENSTEINENSIA

Biblioteca Nacional de Liechtenstein: <http://www.landesbibliothek.li/>

eLiechtensteinensia: <http://www.eliechtensteinensia.li/>

A biblioteca digital da Biblioteca Nacional de Liechtenstein chama-se eLiechtensteinensia. Nela estão guardados livros, jornais e revistas. A digitalização dos jornais e revistas foi iniciada em 2009, sendo realizada por uma empresa externa. No ano seguinte, a coleção foi disponibilizada na “Web”. A página inicial é apresentada na Figura 43, onde se pode visualizar também o mecanismo de busca.

Figura 43 – Página de busca da eLiechtensteinensia

Suchen

Suche im Volltext:

Erweiterte Optionen: keine Suchbegriffe als Wortteil mit Ähnlichkeit Wortstammsuche aktivieren

Von: 1863 2005

Suchen in: Liechtensteiner Vaterland und Vorgänger Liechtensteiner Volksblatt Verschiedene historische Zeitungen

Blättern

Mo	Di	Mi	Do	Fr	Sa	So
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

Liechtensteiner Vaterland und Vorgängerzeitungen

Oberrheinische Nachrichten: [25.4.1914](#) – [30.8.1924](#)

Liechtensteiner Nachrichten: [3.9.1924](#) – [28.12.1935](#)

Liechtensteiner Heimatdienst: [14.10.1933](#) – [28.12.1935](#)

Liechtensteiner Vaterland: [1.1.1936](#) – [31.12.2005](#)

Liechtensteiner Volksblatt

Liechtensteiner Volksblatt: [16.8.1878](#) – [31.12.2005](#)

Verschiedene historische Zeitungen

Liechtensteiner Landeszeitung: [12.4.1863](#) – [7.12.1867](#)

Liechtensteinische Wochenzeitung: [24.1.1873](#) – [28.12.1877](#)

Der Umbruch: [5.10.1940](#) – [12.2.1944](#)

Fonte: Biblio. Nac. Liechtenstein. Disponível em: <<http://www.landesbibliothek.li>>. Acesso em: 13 jan. 2016.

Mesmo se tratando de apenas oito títulos, os jornais compreendem o período de 1863 a 2005, reunindo mais de 500 mil páginas. As publicações disponíveis pertencem ao acervo da BN. Inicialmente, o material foi microfilmado, sendo nos últimos anos também digitalizado e submetido à leitura por OCR.

Apenas um funcionário trabalha diretamente com a coleção na BN, realizando a catalogação dos materiais por título e data. É utilizado o MARC21. Nota-se que alguns artigos apresentam assuntos vinculados, mas isso não ocorre com todos os resultados.

Historiadores e outros pesquisadores são o público mais frequente. Há, em média, 1.000 visitantes por mês. A biblioteca digital oferece dicas aos usuários sobre como melhorar a busca no site <http://www.eliechtensteinensia.li/static/help2.html>, ensinando a utilizar os operadores booleanos e a refinar a busca por meio do uso de aspas ou de asterisco.

O mecanismo de busca permite que se instalem filtros por período, publicação e título (a equipe da eLiechtensteinensia deu a cada página um título, colocando o nome do jornal, a data e a página), e, ainda, que se opte por buscar palavras semelhantes ou palavras iniciadas por determinada palavra ou conjunto de letras. Os resultados são apresentados em uma lista de títulos. Ao se escolher um deles, tem-se acesso a uma página em PDF, que pode ser salva no computador.

6.10 MÉXICO: HEMEROTECA NACIONAL DIGITAL DE MÉXICO

Biblioteca Nacional do México: <http://bnm.unam.mx>

Hemeroteca Nacional Digital de México: <http://www.hndm.unam.mx/index.php/es/>

O acesso à Hemeroteca Nacional dá-se facilmente tanto pela página principal da Biblioteca Nacional mexicana (<http://bnm.unam.mx/>) quanto por meio de um endereço eletrônico próprio – <http://www.hndm.unam.mx/>. A coleção hemerográfica digital compreendia, em janeiro de 2016, nove milhões de páginas, em 947 títulos, do período entre 1722 e 2010.

Como afirmam Curriel Defossé e Jiménez Rivera (2012, p. 2), os responsáveis pela coleção de jornais e revistas da Biblioteca Nacional do México, entidade ligada à Universidad Nacional Autónoma de México (UNAM), sempre estiveram atentos à preservação da coleção, utilizando os recursos que melhor atendiam a esta necessidade em cada época. Uma história curiosa data de 1884, quando foi proposto que houvesse gatos no prédio onde era guardado o acervo para que este ficasse mais protegido dos ratos que ameaçavam roer as obras. Ao longo de décadas, foram desenvolvidos trabalhos de catalogação e encadernação, assim como se buscou melhores condições de armazenamento e resguardo da coleção.

A microfilmagem do acervo, de acordo com os autores, foi iniciada nos anos 1960, sendo a técnica mais segura no período. Tinha início um programa de microfilmagem mantido até hoje, que inclui a compra de materiais microfilmados de outros acervos mexicanos. Na coleção da BN mexicana, há mais de 28 milhões de páginas de jornais e revistas microfilmadas. Nos anos 1990, criou-se um programa complementar voltado à duplicação do material e foram compradas máquinas para a leitura e impressoras para a reprodução dos microfilmes.

Os anos 1990 trouxeram novidades com o avanço das tecnologias de informação – correio eletrônico, página em HTTP, digitalização de livros, acesso do público à rede mundial de computadores, para citar algumas. De repente, o microfilme passou a ser visto como um suporte obsoleto, mas logo se percebeu que se podia escanear o conteúdo ali armazenado de forma muito mais prática. Além disso, com a rapidez com que as tecnologias foram sendo substituídas por outras ainda mais novas, o microfilme mostrou continuar a ser uma ótima forma de preservação de acervos.

Nos anos 2000, a direção da hemeroteca da BN mexicana fez um levantamento entre os prestadores de serviços de tecnologia a fim de encontrar uma empresa apta a digitalizar imagens de jornais e revistas em alta resolução, aplicar o OCR, gerar arquivos HTML e XML,

mecanismos de busca e imagens em que aparecerem as palavras buscadas em destaque. Os resultados foram desalentadores. As empresas sequer tinham *scanners* para microfimes.

A solução apresentou-se ao realizar um contato com a empresa jornalística de Guadalajara *El informador*. Ali não só entendiam do sistema buscado, como já estavam em processo de desenvolvimento de um sistema de ponta, como relatam Curriel Defossé e Jiménez Rivera (2012, p. 8). Estabeleceu-se então uma parceria entre a BN e a empresa. Em 14 de junho de 2002 se iniciaram os trabalhos de “transformação” do acervo e a construção da primeira versão da hemeroteca digital.

A terceira versão foi lançada internamente no ar em 2007, depois de acatar várias sugestões de usuários – de outubro de 2005 a dezembro de 2007, 1752 registraram-se na página da hemeroteca, abrindo 82.774 sessões e realizando 142.838 consultas ao acervo. Durante esse período, levantou-se o perfil dos usuários. Quase 60% do público realizavam uma graduação, sendo oriundos das Ciências Sociais (45%), Humanidades (37%) e Engenharia (6,76%). Entre as razões de acesso estavam: 31% para trabalhos de pesquisa, 19% para trabalhos escolares, 14% para produção de dissertação ou tese e 20% para obter conhecimento de cultura em geral. O perfil se mantém semelhante até os dias de hoje, mostrando que a hemeroteca é acessada por acadêmicos de diferentes níveis e áreas do conhecimento, assim como com propósitos variados.

A versão global da página da hemeroteca foi colocada no ar em 2011, contendo acervo já em domínio público e publicações oficiais dos governos estaduais e federal.

O projeto de digitalização do acervo hemerográfico da Biblioteca Nacional mexicana representou a ascensão ao espaço virtual do mais completo patrimônio do gênero no país. O acervo é composto por jornais que guardam informações sobre acontecimentos políticos, sociais, econômicos e culturais, assim como o desenvolvimento da indústria, da arte, da ciência, da educação etc. O projeto inspirou-se em experiências de outras Bibliotecas Nacionais, como as da Austrália, da Nova Zelândia, do Reino Unido e dos países escandinavos.

Com a criação da hemeroteca digital, amenizaram-se problemas como a manipulação excessiva das publicações, que levava a biblioteca a restringir o empréstimo, a consulta e a reprodução, especialmente da coleção anterior a 1917.

Na página principal, há um menu composto por três itens: busca básica (*Búsqueda básica*), busca avançada (*Búsqueda avanzada*) e ferramentas de consulta (*Herramientas de consulta*). Há ainda um campo em que se pode ler *Búsqueda por palabra...* para uma consulta imediata. A página inicial, remodelada no final de 2015, contém ícones que direcionam o

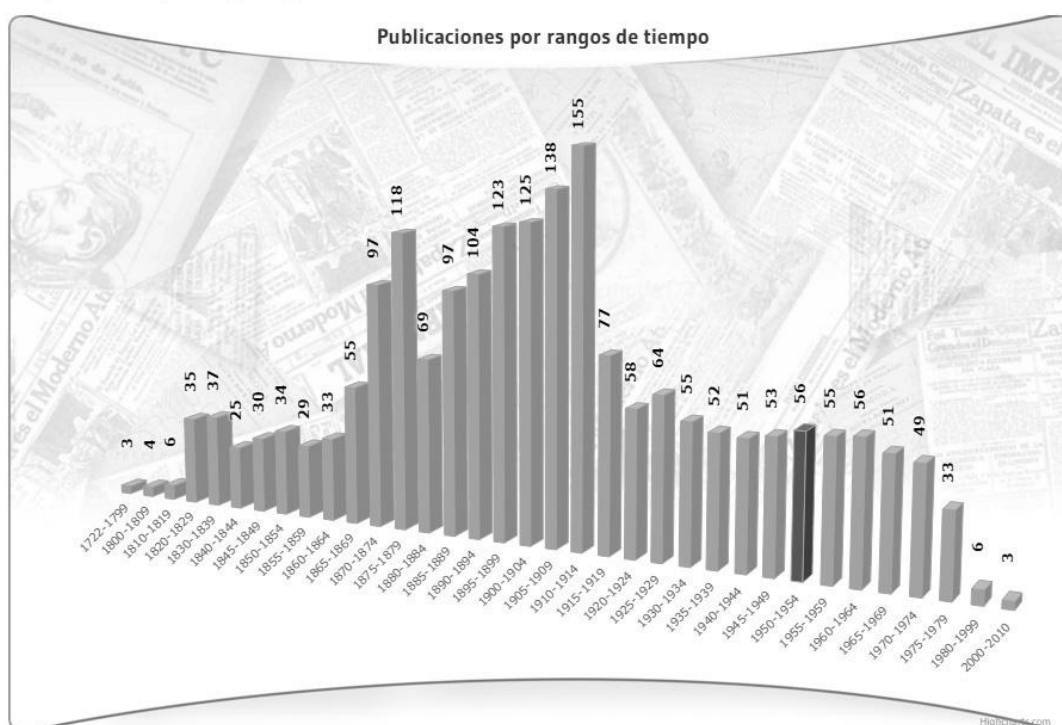
usuário para a página geral da Biblioteca Nacional, um guia de auxílio à navegação, uma página em que são apresentadas as regras de usos das imagens e um link para redes sociais.

Ao escolher a busca básica, o usuário é remetido a uma página na qual se visualiza três abas – por períodos de tempo (Figura 44), por título (Figura 45) e por localização geográfica (Figura 46). Na busca por período de tempo, o consulente escolhe a data a partir de um gráfico, basta clicar na época desejada.

Na busca por título, a partir das letras do alfabeto e dos filtros apresentados em um menu lateral, no qual se pode selecionar país, estado, cidade, tipo de acesso (restrito ou público), idioma, periodicidade, ano e nome da publicação. Nesta opção não é possível inserir palavras-chave. Os resultados são apresentados em forma de tabela ou ficha.

Figura 44 – Página de busca básica da hemeroteca mexicana – por ano

Seleccione un periodo de tiempo en la siguiente gráfica:



Fonte: Biblioteca Nacional do México. Disponível em: <<http://bnm.unam.mx>>. Acesso em: 13 jan. 2016.

Figura 45 – Página de busca básica da hemeroteca mexicana – por título

UNAM Universidad Nacional Autónoma de México

h n d m
HEMEROTECA NACIONAL
DIGITAL DE MÉXICO

iiB
Biblioteca Nacional
Hemeroteca Nacional

Búsqueda básica Búsqueda avanzada Herramientas de consulta Búsqueda por palabra...

Por Título Por Año Geográfica

Seleccione una letra del alfabeto y el sistema le mostrará los títulos correspondientes, así como la posibilidad de consultar todas las páginas del título que más le interese.

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z Todos

PRECISIÓN APLICADA:

PAÍS:
 España (2)
Remover

PRECISAR RESULTADOS:

PAÍS
 México (97)

ESTADO
 Barcelona (1)
 Madrid (1)

CIUDAD
 Barcelona (1)
 Madrid (1)

TIPO DE ACCESO
 Restringido (1)
 Público (1)

IDIOMA
 Español (2)

PERIODICIDAD
 Quincenal (1)
 Semanario (1)

NOMBRE DE LA PUBLICACIÓN
 Abeja: revista científica y literaria ilustrada, La (1)
 Alrededor del Mundo (1)

FECHA
 1866 (1)
 1911 (1)
 1913 (1)

MI BÚSQUEDA: títulos que comienzan con la letra A

Resultados obtenidos: 1 a 2 de 2

Ordenar por: Título (z-a) Desplegar formato: Ficha Resultados por página: 20 Ver

La publicación puede consultarse via Internet
Para consulta exclusivamente dentro de las instalaciones de la HNM

1

Título: Alrededor del Mundo
Publicado entre: 1911-01-04 - 1913-12-07
País: España
Estado: Madrid
Ciudad: Madrid
Periodicidad: Semanario
Idioma: Español
Imágenes digitalizadas: 1429
Tipo de acceso: Restringido

2

Título: Abeja: revista científica y literaria ilustrada, La
Publicado entre: 1866-01-01 - 1866-01-01
País: España
Estado: Barcelona
Ciudad: Barcelona
Periodicidad: Quincenal
Idioma: Español
Imágenes digitalizadas: 869
Tipo de acceso: Público

Resultados obtenidos: 1 a 2 de 2

Fonte: Biblioteca Nacional do México. Disponível em: <<http://bnm.unam.mx>>. Acesso em: 13 jan. 2016.

Figura 46 – Página de busca básica da hemeroteca mexicana – por localização



Fonte: Biblioteca Nacional do México. Disponível em: <<http://bnm.unam.mx>>. Acesso em: 13 jan. 2016.

Por meio da busca avançada, o usuário pode buscar por título e conteúdo com auxílio de operadores apresentados como “todas estas palavras”, “esta frase exata” ou “ao menos esta palavra”. Pode-se ainda escolher o título da publicação a partir de uma lista, o país, o estado e a cidade, data, idioma, periodicidade e tipo de acesso. Os resultados são apresentados como se os jornais estivessem expostos em uma estante. Ao lado da imagem da página, há informações detalhadas sobre o documento.

Há ainda a aba Ferramentas de busca, onde se tem acesso aos catálogos de publicações periódicas da Biblioteca Nacional do México.

6.11 NOVA ZELÂNDIA: PAPERS PAST

Biblioteca Nacional da Nova Zelândia: <http://natlib.govt.nz>

Papers Past: <http://paperspast.natlib.govt.nz/cgi-bin/paperspast>

Chama-se Papers Past a hemeroteca digital da Biblioteca Nacional da Nova Zelândia. Desde 2001, quando começou a ser desenvolvido, o processo de apresentação dos objetos

digitalizados já passou por três grandes alterações. Inicialmente, era criado apenas um PDF das edições de jornais que eram inseridas em uma base de dados navegável.

Em um segundo momento, houve uma reestruturação sistemática de processos, padrões e entrega de informações. Esta fase incluiu um fluxo de trabalho de digitalização em grande escala – quando foram inseridos os microfimes no sistema e passou-se a utilizar TIFF, GIF e PNG, além da extração de termos via OCR para utilização de metadados METS-ALTO XML. “Isto permitiu aos usuários navegarem e realizarem buscas na coleção, assim como permitiu aumentar a digitalização dramaticamente”, afirma Emerson Vandy, gerente de serviços digitais do Papers Past.

A terceira fase está atualmente em desenvolvimento, mas segue as mesmas orientações da segunda etapa, crescendo, porém, em extensão e escala. O Papers Past já continha, em setembro de 2015, quatro milhões de páginas de 120 jornais veiculados entre 1839 e 1948. Há algumas publicações que vão até 1975. Não há material mais recente por questões de direitos autorais. O foco da coleção é dar acesso a jornais antigos.

Vandy estima que 10% dos jornais publicados na Nova Zelândia tenham sido digitalizados. Ele explica que o objetivo não é colocar *on-line* toda a coleção, até por limitações de orçamento, mas priorizar publicações que atendam às necessidades de pesquisadores em termos de cobertura geográfica, temporal e populacional. Atualmente, trabalham duas pessoas no projeto, sendo uma delas bibliotecária. De acordo com dados do *Google Analytics*, os usuários têm em média 30 anos, e 54% são homens. Mensalmente, entre 100 mil e 150 mil visitam o Papers Past (Figura 47).

Não há indexação de conteúdos dos jornais postos *on-line*. Apenas é realizada uma categorização do conteúdo – artigos, ilustrações, anúncios. Antes de colocar no ar as páginas, elas são conferidas, para ver se são capazes de gerar leitura pelo OCR. Quando a qualidade é muito baixa, há uma política de não colocá-las na rede.

Os jornais são inseridos no catálogo da Biblioteca Nacional da Nova Zelândia por título, título alternativo, código, região, ano inicial e final da coleção, além da informação se a publicação continua ativa ou se já foi descontinuada. É utilizado o formato MARC21 na catalogação.

Tão logo acessa a hemeroteca, o usuário já se depara com o mecanismo de busca, no qual a consulta poderá ser realizada por palavras-chave (*searching*) ou por data, região e título (*browsing*), resultando em uma lista que levará à imagem do jornal, que pode ser salva em GIF e PDF.

Figura 47 – Página inicial do *Papers Past*

Fonte: National Library of New Zealand. Disponível em: <<http://natlib.govt.nz>>. Acesso em: 15 jan. 2016.

Na busca avançada (Figura 48), é possível selecionar um ou mais periódicos específicos, tipo de conteúdo (artigos, ilustrações ou propagandas) e período.

Esta hemeroteca tem um diferencial: o material está armazenado por assunto, página, assim como pelo tipo de material – reportagens, propagandas e ilustrações. Existe ainda um setor que atende às necessidades dos usuários que têm dificuldade em realizar uma busca.

Os resultados (Figura 49) são apresentados por *best match first* (melhor resultado primeiro), data, título do artigo, título do periódico e tipo de material (representados por diferentes ícones).

Figura 48 – Resultados da busca no *Papers Past*

Fonte: National Library of New Zealand. Disponível em: <<http://natlib.govt.nz>>. Acesso em: 15 jan. 2016.

Figura 49 – Visualização dos resultados da busca no *Papers Past*

Papers Past Home Introduction Search Browse

Papers Past > New Zealand Herald > 27 July 1940 > Page 1 > Glimpse of Rio de Janeiro's business area

Glimpse of Rio de Janeiro's business area
New Zealand Herald, Volume LXXVII, Issue 23719, 27 July 1940, Page 1

► About this newspaper View computer-generated text



Glimpse of Rio de Janeiro's business area

This article has been automatically clipped from the New Zealand Herald, organised into a single column, then optimised for display on your computer screen.

This newspaper was digitised in partnership with Auckland Libraries and APN.

Te Puna Mātauranga o Aotearoa
NATIONAL LIBRARY OF NEW ZEALAND
About this site | Site map | Accessibility | Contact us | Terms of use | Copyright guide

1940-07-27

Fonte: National Library of New Zealand. Disponível em: <<http://natlib.govt.nz>>. Acesso em: 15 jan. 2016.

Como em outras hemerotecas estudadas, a *Papers Past* oferece a possibilidade de se copiar o conteúdo em forma de texto, detectado por OCR.

6.12 REINO UNIDO: BRITISH NEWSPAPER ARCHIVE

British Library: <http://www.bl.uk>

British Newspaper Archive: <http://www.britishnewspaperarchive.co.uk/>

A coleção do *British Newspaper Archive* (BNA, Figura 50) reúne jornais da Inglaterra, Irlanda e Escócia. As edições cobrem especialmente o século XIX, mas há algumas publicações do século XVIII. O acervo foi iniciado pela biblioteca do Museu Britânico (*British Museum*) e, desde 1972, assumido e continuado pela Biblioteca Nacional britânica (*British Library*), que sistematicamente coleta jornais e revistas por meio da lei do depósito legal, assim como pela compra e troca de exemplares, segundo descreve King (2005, p. 166).

Uma das primeiras medidas para a preservação do acervo foi a microfilmagem, uma vez que a qualidade e a durabilidade do filme é bastante conhecida e confiável. Os negativos de microfilmes permitem a reprodução de cópias, além de durarem o tempo estimado de 200 anos. Nos processos de digitalização de acervos de Bibliotecas Nacionais, os microfilmes foram os primeiros a serem escaneados.

Atualmente, a coleção de jornais e revistas é publicada e gerenciada pela *DC Thomson Family History* em colaboração com a Biblioteca Nacional britânica. Por meio da parceria, com duração de 10 anos (iniciada em 2011), pretende-se digitalizar 40 mil páginas de jornal. Diferente de outras Bibliotecas Nacionais, a britânica não obteve recursos públicos para criar a hemeroteca digital. A alternativa foi buscar um parceiro externo privado.

Figura 50 – Página inicial do *The British Newspaper Archive*

The screenshot shows the homepage of The British Newspaper Archive. At the top, there are navigation links for 'Sign in', 'Subscribe', and 'Register'. Below this is a main navigation bar with links for 'HOME', 'BLOG', 'HELP & ADVICE', 'ABOUT', 'REGISTER', and 'SEARCH'. The central banner features the text 'Explore over 200 years of history' and 'Millions of stories: from your street to the world', accompanied by a search bar and a 'SEARCH' button. Below the banner, there are three main sections: 'Browse by newspaper title' with a dropdown menu, 'Browse by place of publication' with a map of the United Kingdom, and 'Browse by date range' with a bar chart showing the distribution of titles over time. The page also displays statistics: '12,671,779 pages: see latest additions' and '560 titles online'. At the bottom, there is a footer with a Twitter link and a location indicator 'London, England'.

Fonte: British Library. Disponível em: <<http://www.bl.uk>>. Acesso em: 14 jan. 2016.

O protótipo de um sistema de digitalização, indexação e preservação de jornais antigos da Biblioteca Nacional britânica começou a ser pensado na primeira metade de 2001, utilizando-se jornais livres de direitos autorais. De acordo com King (2005, p. 174), o objetivo principal do projeto era permitir o acesso *on-line* aos jornais antigos, especialmente aqueles do século XIX, mas também se esperava:

- Digitalizar uma massa crítica de jornais microfilmados em um tempo relativamente curto de tempo;
 - Indexar automaticamente os documentos digitalizados;
 - Entregar os documentos em um formato *on-line* sofisticado;
 - Prover uma recuperação avançada dos textos; e
 - Apresentar os resultados para o maior número possível de pessoas.
- (KING, 2005, p. 174)

A digitalização de parte do acervo foi uma resposta a alguns desafios encontrados pelos administradores da coleção: jornais que se deterioravam rapidamente devido à péssima qualidade do material; o local de armazenamento em Colindale (onde inicialmente ficavam os jornais) estava completamente cheio e era muito caro modificá-lo; as condições de armazenamento eram inapropriadas; o uso pelo público estava deteriorando ainda mais o estado dos jornais; 15% da coleção não podia mais ser usada, pois estava muito frágil; outros 19% da coleção estavam quase chegando ao estado de não poderem mais ser emprestados; os serviços oferecidos em Colindale não atendiam às necessidades dos pesquisadores.

Inicialmente foram registrados muitos problemas inerentes à natureza de jornais antigos, o que se revelou um desafio a mais para os envolvidos no projeto, como descreve King (2005, p. 175). O leiaute das publicações era mais complexo do que o planejado, havia mudanças radicais de diagramação de tempos em tempos, os títulos tinham letras grandes demais, os espaços entre as linhas eram estreitos, a qualidade da imagem microfilmada era ruim, a curvatura dos volumes encadernados dificultava a digitalização “plana” dos jornais, havia sujeira em algumas páginas e incomodava o desalinhamento nas linhas dos artigos e dos anúncios.

O processo de digitalização foi dividido em duas partes: escaneamento de precisão e processamento de imagem, o que gerou um arquivo para cada imagem captada. O sistema escolhido à época foi o *Olive Software System*, que implantava componentes *ActivePaper Archive*. Já nesta primeira experiência foram usadas palavras geradas pelo OCR, que foram guardadas em formato XML (*Extensible Markup Language*).

As buscas no acervo do BNA são gratuitas, mas o acesso aos dados recuperados é cobrado. Ao se registrar, o usuário ganha o acesso a três resultados. Existem dois tipos de pacotes pagos de acesso: o primeiro custa £12.95 com duração de 30 dias e o segundo, £79.95

com direito a um ano de uso.

O acesso ao acervo é gratuito quando realizado nos salões de leitura da biblioteca, em St. Pancras, Londres, e em Boston Spa – neste segundo endereço está armazenada desde 2014 a coleção física de jornais. De acordo com uma regra criada pela Biblioteca Nacional, os usuários recebem cópias em microfilme ou digitalizadas para consulta.

Quando o material já se encontra copiado, seja em formato digital ou microfilme, o documento original é etiquetado como “restrito” e arquivado no depósito de Boston Spa. O prédio que abriga a coleção chama-se *National Newspaper Building* (Figura 51). O armazém onde ficam os jornais e revistas tem nível de oxigênio baixo para evitar a deterioração dos itens e o transporte das publicações é totalmente automatizado.

Figura 51 – *National Newspaper Building* do BNA, em Boston Spa



Fonte: British Library. Disponível em: <<http://goo.gl/Xy0c2s>>. Acesso em: 14 jan. 2016.

Em janeiro de 2016, o BNA apresentava-se composto por 12,6 milhões de páginas de 560 jornais de todas as partes do Reino Unido e da Irlanda. Stephens, Brazier e Spence (2014, p. 206) estimam que a coleção da Biblioteca Nacional britânica contenha 750 milhões de páginas, de jornais de três séculos. As publicações veiculadas antes de 1900 estão livres de direitos autorais. Para digitalizar muitos dos demais anos, especialmente os do século XX, a *DC Thompson Family History* obteve a permissão junto às editoras responsáveis.

As buscas podem ser feitas com palavras-chave, em um modelo semelhante ao do *Google*, sendo oferecida apenas uma caixa para serem digitados os temas ou frases. Há também possibilidade de busca avançada, na qual se pode combinar data, região, título, tipo

de artigo e etiquetas que outros usuários tenham acrescentado a obras do acervo. O usuário conta ainda com um menu para refinamento da busca e uma página com dicas sobre como melhorar a busca: <http://www.britishnewspaperarchive.co.uk/content/help>.

O BNA utiliza a *fuzzy search*, tipo de busca que primeiro recupera arquivos que contenham a palavra exatamente como ela foi digitada (ex.: *archive*), para logo em seguida recuperar variantes do termo (ex.: *archived*, *archives*, *archiving* etc.), o que aumenta a revocação. O usuário pode optar por não utilizar este recurso.

Os resultados são apresentados em forma de capas de jornais (Figura 52). Para visualizar, após ter escolhido um tipo de pacote de acesso, basta clicar sobre o *thumbnail* desejado. O material escolhido pode ser baixado ou impresso. Os usuários registrados podem acrescentar *tags*. Como em outras hemerotecas analisadas, o BNA utiliza-se de OCR, permitindo o acesso a uma versão em formato de texto.

Figura 52 – Resultados da busca no BNA

The screenshot shows the search results for 'rio de janeiro' on the British Newspaper Archive website. The search bar at the top contains 'rio de janeiro' and a 'Search' button. Below the search bar, there are options for 'Keep filters' and 'Exact search'. The results are displayed in a list format, with each result showing a thumbnail of a newspaper page, the title 'RIO DE JANEIRO', a quote of the article's content, the date, the newspaper name, and the location. The first result is from 'Saturday 25 April 1936, Sheffield Independent, South Yorkshire, England'. The second result is from 'Tuesday 07 May 1861, Belfast Mercury, Antrim, Northern Ireland'. The third result is from 'Monday 05 March 1900, Western Times, Devon, England'. On the left side of the page, there are filter options for 'Filter by Date' and 'Filter by Region', both showing the number of articles for each category.

Fonte: British Library. Disponível em: <<http://www.bl.uk>>. Acesso em: 14 jan. 2016.

6.13 RESUMO DAS CARACTERÍSTICAS DAS HEMEROTECAS ESTRANGEIRAS

A fim de se ter uma visão geral das hemerotecas estrangeiras pesquisadas, neste subcapítulo são apresentadas algumas de suas principais características, como ano de lançamento, tamanho e cobertura, além de particularidades dos mecanismos de busca.

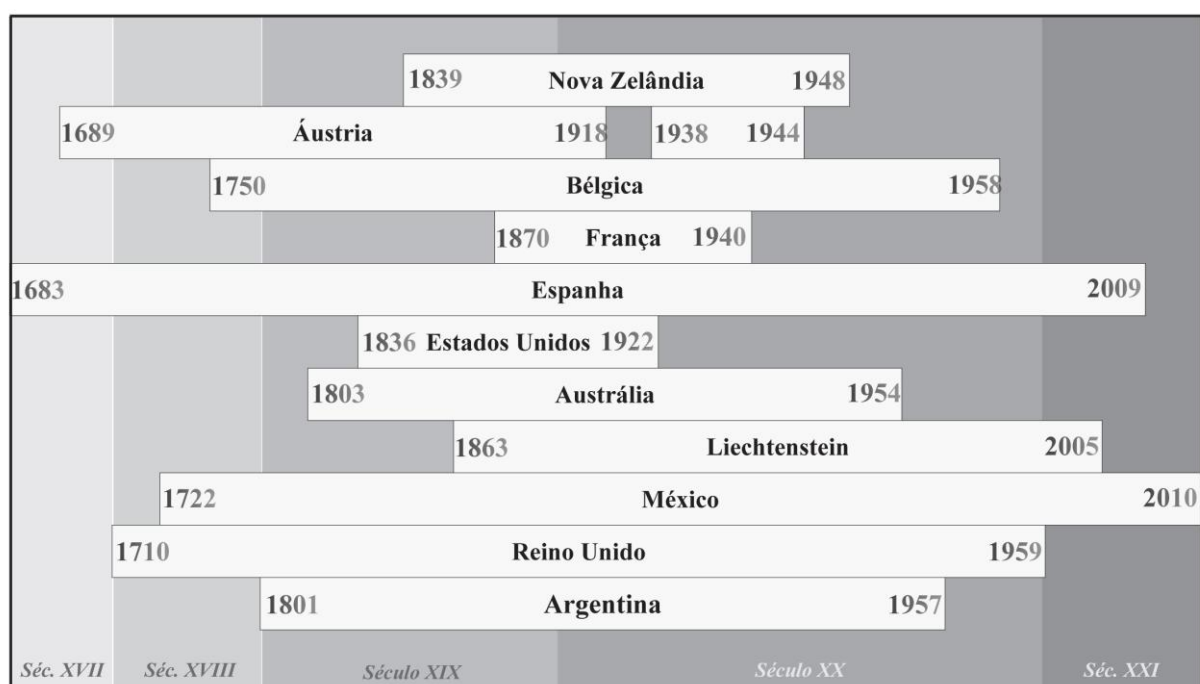
Entre as hemerotecas analisadas (Tabela 1), a da Biblioteca Nacional da Nova Zelândia é a mais antiga na “Web”, tendo sido lançada em 2001. Nesse mesmo ano, a Biblioteca Nacional do Reino Unido começava a digitalizar microfimes, mas a versão *on-line* do sistema só estaria no ar 10 anos mais tarde. Apesar de ter lançado a hemeroteca digital em 2007, a Espanha foi a que começou a desenvolver o trabalho de digitalização mais cedo, em 1997.

Tabela 1 – Data de início do projeto e de lançamento das hemerotecas digitais

Hemeroteca	Ano de lançamento	Ano de início do projeto
Nova Zelândia	2001	-
Áustria	2003	2003
Bélgica	2004	2004
França	2006	-
Espanha	2007	1997
Estados Unidos	2007	-
Austrália	2009	2007
Costa Rica	2009	2009
Liechtenstein	2009	-
México	2011	2002
Reino Unido	2011	2001
Argentina	2014	2011

Entre as hemerotecas, a da Biblioteca Nacional da Espanha é a que apresenta o maior período de cobertura, indo de 1683 a 2009. Com exceção desta e das hemerotecas de Liechtenstein e do México, as demais oferecem acesso *on-line* apenas a jornais que já caíram em domínio público ou àqueles autorizados pelas empresas que os editam. O século com maior cobertura é o século XIX (conforme evidenciado pela análise comparativa apresentada na Figura 53), compreendendo aproximadamente 70 anos de cobertura, de 1801 a 1870.

Figura 53 – Período de cobertura das hemerotecas



Muitas Bibliotecas Nacionais têm acervos digitalizados muito maiores do que o disponibilizado na “Web”. Exemplo disso é a Biblioteca Real da Bélgica, que conta com 3,7 milhões de páginas digitalizadas, mas oferece acesso *on-line* a apenas pouco mais de 1,2 milhão. O restante da coleção digitalizada por ser consultada somente *in loco*.

O maior acervo de jornais e revistas acessível na “Web” é o da Biblioteca Nacional da Austrália. Na seção de jornais do Trove, o usuário tem acesso a mais de 18 milhões de páginas em mil títulos. Em seguida, com base nas informações coletadas junto às BN no período entre julho de 2015 e janeiro de 2016, vêm a coleção da Espanha (17 mi. em 1.275 títulos), da Áustria (14 mi. em 600 títulos) e do Reino Unido (12,6 mi. em 560 títulos), como detalhado na Tabela 2.

Tabela 2 – Número de títulos e de páginas digitalizados

Hemeroteca	Número de títulos	Número de páginas
Costa Rica	518	-
Estados Unidos	1.800	-
Argentina	8.000	-
Liechtenstein	8	500 mil
Bélgica	-	1,2 milhão
França	-	1,5 milhão
Nova Zelândia	120	4 milhões
México	947	9 milhões
Reino Unido	560	12,6 milhões
Áustria	600	14 milhões
Espanha	1.275	17,8 milhões
Austrália	1.000	18 milhões

O recurso de reconhecimento óptico de caracteres (OCR) é amplamente usado nas coleções digitalizadas das Bibliotecas Nacionais pesquisadas. Das hemerotecas consultadas apenas Argentina e Costa Rica não aplicam OCR. Nesses casos, o usuário insere palavras-chave no buscador, mas só obterá algum resultado se a palavra estiver contida no nome do jornal. A busca não engloba o conteúdo dos jornais e revistas. No Quadro 8 são apresentados os parâmetros/metadados disponíveis em cada hemeroteca estudada.

No caso da Costa Rica, o usuário tem mais opções para refinar a busca, sendo possível escolher assunto, local e período. A hemeroteca digital da Argentina apenas permite ordenar o acervo baseada em local, data e nome da publicação. Como se trata da hemeroteca instalada somente em 2015, ou seja, recentemente, estima-se que venha a realizar melhorias na forma de recuperar o conteúdo. Na primeira pesquisa realizada na “Web” para localizar hemerotecas, em 2014, a Biblioteca Nacional da Argentina não figurava entre os países com hemerotecas digitais. No ano seguinte, apresenta novidades com o recém-lançado Trapalanda. Já nas demais hemerotecas analisadas é possível pesquisar diretamente no conteúdo das publicações. O *software* OCR permite o reconhecimento das palavras, levando o usuário às páginas em que ocorrem as palavras-chave inseridas na busca.


Quadro 8 – Parâmetros/metadados usados nas buscas simples e avançada

Hemeroteca Digital	Busca simples	Busca avançada
Argentina	Palavra-chave (título apenas)	<i>Não oferece</i>
Austrália	Palavra-chave (conteúdo)	Título, data, local, idioma
Áustria	Palavra-chave (conteúdo)	Título, data, local, idioma, assunto
Bélgica	Palavra-chave (conteúdo)	Título, data, local, idioma
Costa Rica	Palavra-chave (título apenas)	Título, ano, dia, mês e assunto
Espanha	Palavra-chave (conteúdo)	Título, local, data
Estados Unidos	Palavra-chave (conteúdo), local, período	Título, local, período, idioma, data,
França	Palavra-chave (conteúdo)	Data, título, língua, tipo de documento, assunto, biblioteca
Liechtenstein	Palavra-chave (conteúdo)	Data, título
México	Título, local, data	Título, local, data, idioma, periodicidade e tipo de acesso
Nova Zelândia	Palavra-chave (conteúdo)	Título, local, data, tipo
Reino Unido	Palavra-chave (conteúdo)	Título, local, data, tipo, tags

As interfaces oferecidas pelas hemerotecas são bastante variadas, como se observa na compilação apresentada a seguir (Figura 54). Na mesma página em que se realiza a busca simples, por palavra-chave, já se pode refinar a busca por título, data e local, como nas páginas da Austrália, Bélgica, Espanha, Nova Zelândia e Reino Unido. Na página norte-americana do Chronicling America, pode-se escolher o estado e o período. Já na página de busca da hemeroteca da Biblioteca Nacional de Liechtenstein estão listados os oito títulos que compõem a coleção e pode-se escolher a data de interesse.

Figura 54 – Leiaute das páginas de busca das hemerotecas estudadas

Austrália


Trove Digitised newspapers and more
[Home](#) [About](#) [Browse](#) [Help](#)

Find an article

Advanced Search

Find an issue

by Title


1. The Sydney Morning Herald
2. The Argus
3. Northern Territory Times and Gazette
4. Brisbane Courier
5. The Australian Women's Weekly

by Date

1803 ▼

JAN	FEB	MAR	APR			
MAY	JUN	JUL	AUG			
SEP	OCT	NOV	DEC			
S	M	T	W	T	F	S

by State

NT		QLD
WA		NSW
SA		ACT
TAS		VIC

Suche:

 Royal Library of Belgium
NL FR EN

[Catalogue of the KBR](#) | [Other catalogues](#) | [Legal deposit](#) | [BelgicaPress](#)
[Contact](#) [Help](#)

Search

text All of these words ▼ + period dd-mm-yyyy to dd-mm-yyyy

newspaper title Select... language All ▼


Browse the newspapers by date

Year... Newspaper...

January	February	March	April
May	June	July	August
September	October	November	December

© 2015 Royal Library of Belgium - [Legal notice](#)

Estados Unidos



LIBRARY OF CONGRESS


ASK A LIBRARIAN

DIGITAL COLLECTIONS

LIBRARY CATALOGS

Search **GO**

The Library of Congress > [Chronicling America](#)



CHRONICLING AMERICA
Historic American Newspapers

Search America's historic newspaper pages from 1836-1922 or use the U.S. Newspaper Directory to find information about American newspapers published between 1690-present. Chronicling America is sponsored jointly by the [National Endowment for the Humanities](#) and the Library of Congress. [Learn more](#) >

Search Pages

Advanced Search

All Digitized Newspapers 1836-1922

US Newspaper Directory, 1690-Present

All states + from 1836 to 1922 + enter one or more search words **GO**


Pages Available: 9,990,550 Print Subscribe Share/Save Give Feedback

- [About Chronicling America](#)
- [About the Site and API](#)
- [Recommended Topics](#)
- [Help](#)


More Resources

- [National Digital Newspaper Program](#)
- [NDNP Award Recipients](#)
- [Newspaper and Current Periodicals Reading Room](#)
- [Ask LC Newspaper & Current Periodicals Librarian](#)
- [Historic Newspapers on Flickr](#) (part of the LC Flickr Commons photostream)


100 Years Ago Today: 05/10/1915 (63 issues)



Hopkinsville Kentuckian. (8pp.)
Hopkinsville, Ky.



The Seattle star. (8pp.)
Seattle, Wash.




The Fairmont West Virginian. (10pp.)
Fairmont, W. Va.

Connect with the Library

All ways to connect >

Find Us On



Subscribe & Comment

RSS & E-Mail

Blogs

Download & Play

Podcasts

Webcasts

iTunesU

Questions

Ask a Librarian

Contact Us

França

(BnF Gallica

Rechercher...

dans tout Gallica

Accueil > Presse et revues > PAR TYPES DE DOCUMENTS > Presse et revues

Rechercher dans tout Gallica

- dans tout Gallica
- dans les livres
- dans les manuscrits
- dans les cartes
- dans les images
- dans la presse et revues
- dans les enregistrements sonores
- dans les partitions
- dans les objets

Les principaux quotidiens

Presse par thématiques >

Presse locale et régionale >

Presse par pays ou zones géographiques >

Journaux de tranchées >

Journaux clandestins de la Résistance >

Revue savantes >

Si l'on considère *La Gazette* de Théophraste Renaudot, on peut dire que c'est à partir de 1777 l'ère de la presse d'information.

1789 instaure la liberté de la presse. Les feuilles éclosent en nombre (*Le Père Duchesne*, *L'Ami du Peuple*) contre la censure agite tout le XIXe siècle qui s'achève par un « âge d'or de la presse » à la faveur de

Dès 1836, les formules novatrices du *Siècle* et de *La Presse* ouvrent la marche de cet essor. Le titre

Liechtenstein

LIECHTENSTEINER LANDESBIBLIOTHEK

eLiechtensteinensia

Suchbegriffe

Suche

Home

Zeitungen

Bücher

Zeitschriften

Volksblatt

Wochenzeitung

Liechtensteiner Landeszeitung

Wochenzeitung

Vaterland

Wochenzeitung

Nachrichten

Wochenzeitung

Liechtensteiner Volksblatt

Wochenzeitung

Hilfe zur Suche in Zeitungen

Suchen

Suche im Volltext: **Suche**

Erweiterte Optionen: keine Suchbegriffe als Wortell mit Ähnlichkeit Wortstammuche aktivieren

Von: Bis:

Suchen in: Liechtensteiner Vaterland und Vorgänger Liechtensteiner Volksblatt Verschiedene historische Zeitungen

Blättern

Jun	1960					
Mo	Di	Mi	Do	Fr	Sa	So
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

Liechtensteiner Vaterland und Vorgängerzeitungen

Oberheirische Nachrichten: 25.4.1914 – 30.8.1924

Liechtensteiner Nachrichten: 3.9.1924 – 28.12.1925

Liechtensteiner Heimatdienst: 14.10.1933 – 28.12.1935

Liechtensteiner Vaterland: 1.1.1936 – 31.12.2005

Liechtensteiner Volksblatt

Liechtensteiner Volksblatt: 16.8.1878 – 31.12.2005

Verschiedene historische Zeitungen

Liechtensteiner Landeszeitung: 12.4.1863 – 7.12.1867

Liechtensteinsche Wochenzeitung: 24.1.1872 – 28.12.1877

Der Umbruch: 5.10.1940 – 12.2.1941

México

Búsqueda básica Búsqueda avanzada Herramientas de consulta Búsqueda por palabra...

Por Título Por Año Geográfica

Seleccione una letra del alfabeto y el sistema le mostrará los títulos correspondientes, así como la posibilidad de consultar todas las páginas del título que más le interese.

A • B • C • D • E • F • G • H • I • J • K • L • M • N • O • P • Q • R • S • T • U • V • W • X • Y • Z • Todos

MI BÚSQUEDA: títulos que comienzan con la letra A

Resultados obtenidos: 1 a 2 de 2

Ordenar por: Título (z-a) desplegar formato: Ficha Resultados por página: 20 Ver

La publicación puede consultarse via Internet
Para consulta exclusivamente dentro de las instalaciones de la HNM

Nova Zelândia

PAPERSPAST English | [Māori](#)

Papers Past Home Introduction Search Browse

New Zealand Gazette
Masthead of New Zealand Tablet
09 July 1853

Papers Past contains more than three million pages of digitised New Zealand newspapers and periodicals. The collection covers the years 1839 to 1948 and includes 120 publications from all regions of New Zealand.

Latest additions to Papers Past (September 2015):

- [Evening Star \(1865-1920\)](#)

Find out more

Search Newspapers

Exact phrase
Any of your words
All of your words

Browse Newspapers

BY DATE: View all newspapers and periodicals by date. [Go](#)

BY REGION: View all newspapers and periodicals by region. [Go](#)

BY TITLE: View all newspapers and periodicals by title. [Go](#)

NATIONAL LIBRARY OF NEW ZEALAND About this site | Site map | Accessibility | Contact us | Terms of use | Copyright guide

Reino Unido

Sign-in Subscribe Register In partnership with **BRITISH LIBRARY**

HOME BLOG HELP & ADVICE ABOUT REGISTER SEARCH

The British Newspaper Archive
One place, millions of stories

Explore over 200 years of history
Millions of stories: from your street to the world

Search for people, places, events and more SEARCH

Show search tips > Advanced search >

Browse by newspaper title Choose a title Browse by place of publication Choose a county

12,671,779 pages: see latest additions

560 titles online

Browse by date range

1710-1759 1760-1809 1810-1859 1860-1909 1910-1959

@Susan1278 Hi Susan. We're having a few issues at the mo which we'll have sorted ASAP. Sorry for the inconvenience

Os resultados recuperados costumam ser apresentados em listas de *links* ou de títulos ou por meio das imagens das páginas em que se encontra a palavra-chave desejada (Quadro 9). Ao clicar no resultado, invariavelmente o usuário é direcionado para uma imagem da página escolhida. Em todas as hemerotecas pesquisadas é possível salvar o resultado em PDF.

A seguir, serão abordadas particularidades das hemerotecas pesquisadas no que diz respeito a apresentação dos resultados.

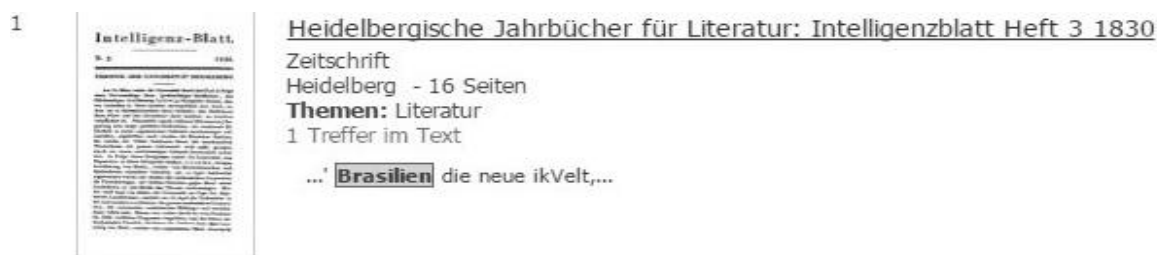
Quadro 9 – Formato apresentação dos resultados

Hemerotecas	Forma predominante
Áustria	Lista de títulos
Austrália	Lista de <i>links</i>
Bélgica	Lista de títulos
Espanha	Imagem das páginas
Estados Unidos	Imagem das páginas
França	Imagem das páginas
Liechtenstein	Lista de <i>links</i>
México	Imagem das páginas e lista de <i>links</i>
Nova Zelândia	Lista de <i>links</i>
Reino Unido	Imagem das páginas

Áustria

Apresenta a imagem da página recuperada (Figura 55), assim como o endereço eletrônico. São listados o nome do jornal, o número de páginas, assim como o assunto principal (no exemplo, literatura) e a quantidade de vezes que ocorre a palavra-chave inserida na busca. O trecho em que aparece a palavra-chave também é mostrado no resultado. Em um menu à esquerda, o usuário pode filtrar os resultados por tipo de veículo (jornal ou revista), local de publicação, idioma, período de cobertura ou assunto.

Figura 55 – Detalhe da apresentação de resultados no ANNO



Fonte: Biblioteca Nacional da Áustria. Disponível em: <<http://anno.onb.ac.at>>. Acesso em 10 fev. 2016.

Austrália

A palavra-chave buscada aparece no título do resultado em forma de *link*, que leva à imagem da página recuperada (Figura 56). Em seguida, lê-se o nome do jornal, o estado em que foi editado, assim como o período de cobertura, a data da página recuperada, o número da página e de que tipo de texto se trata (artigo). O resultado traz ainda o trecho em que a palavra ocorre e o número de palavras. Em alguns dos resultados, há *tags* inseridas pelos usuários, assim como as correções já sugeridas ao texto recuperado por OCR. Em um menu à esquerda, o usuário pode filtrar os resultados por estado (local), título, categoria, se é ou não ilustrado, década e número de palavras.

Figura 56 – Detalhe da apresentação de resultados no Trove



Fonte: National Library of Australia. Disponível em: <<http://trove.nla.gov.au>>. Acesso em: 10 fev. 2016.

Bélgica

Os resultados apresentam-se de maneira simplificada (Figura 57), apenas uma lista dos jornais em que há ocorrência da palavra-chave escolhida e o número de páginas em que aparecem essas ocorrências, apontando quantas estão *on-line* e quantas somente no acervo disponível na sede da Biblioteca Nacional. Ao escolher o jornal de seu interesse, o usuário é levado para uma segunda página, onde há a imagem do primeiro resultado e uma lista com as páginas seguintes. O único filtro possível é escolher entre visualizar somente a coleção *on-line*, a coleção física ou as duas.

Figura 57 – Detalhe da apresentação de resultados no BelgicaPress

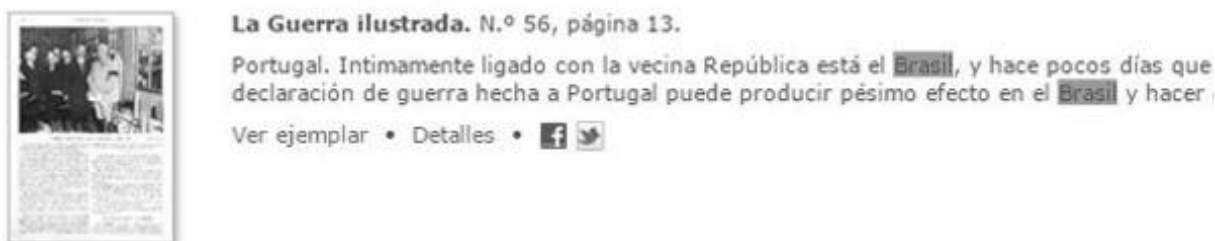


Fonte: KBR. Disponível em: <<http://opac.kbr.be/belgicapress.php?lang=FR>>. Acesso em 10 fev. 2016.

Espanha

O resultado é composto pela imagem da página recuperada (Figura 58), assim como o título, o número da edição, página e o trecho em que aparece a ocorrência. O usuário pode seguir para o exemplar (*ver ejemplar*) ou conferir detalhes (texto recuperado por OCR) sobre o resultado. Há opção de compartilhar no *Facebook* e/ou no *Twitter*. Concluída a busca, não é possível refiná-la.

Figura 58 – Detalhe da apresentação de resultados na hemeroteca espanhola

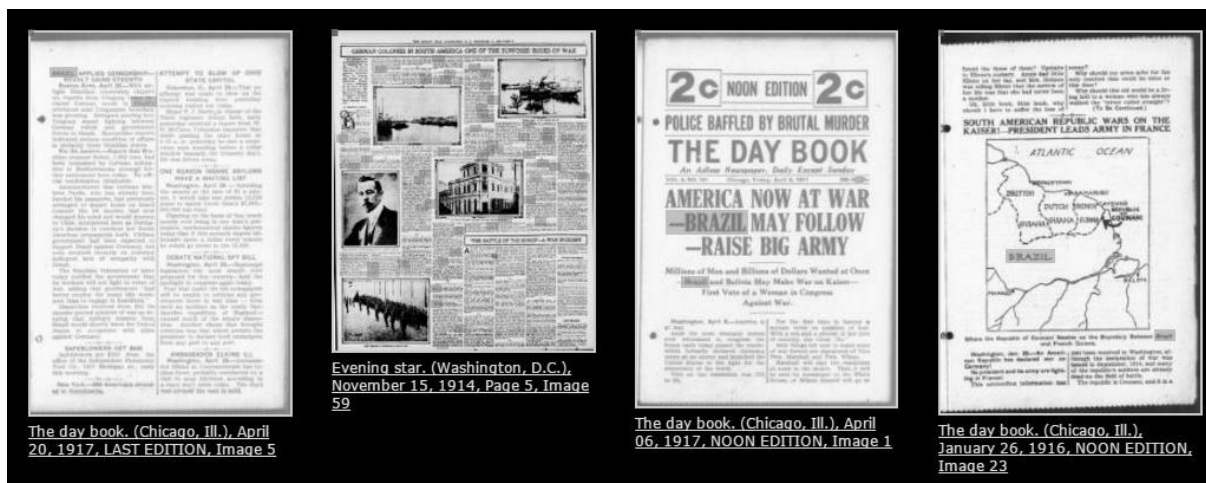


Fonte: Hemeroteca Digital da BNE. Disponível em: <<http://hemerotecadigital.bne.es>>. Acesso em: 10 fev. 2016.

Estados Unidos

Os resultados são apresentados com a imagem da página (Figura 59), o título da publicação, o local, data completa, tipo de edição (matutino, vespertino) e número da imagem. Não é possível filtrar as informações recuperadas. Ao escolher uma das páginas, o usuário é direcionado para a imagem maior, que poderá ser salva no formato preferido (texto, PDF, JPG).

Figura 59 – Detalhe da apresentação de resultados no *Chronicling America*

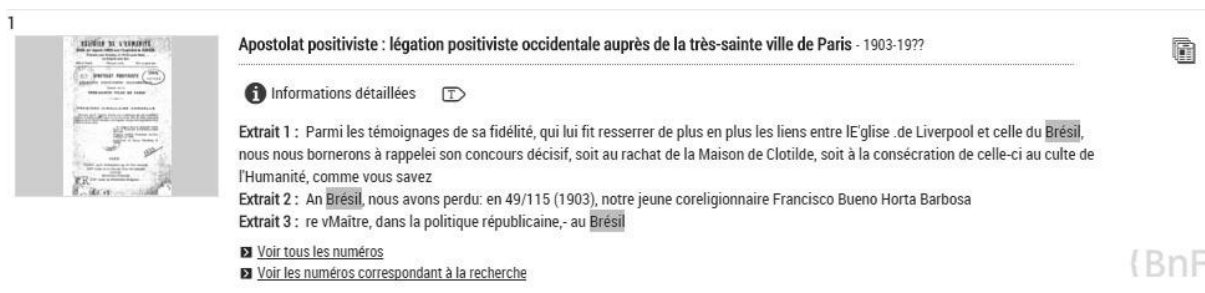


Fonte: Library of Congress. Disponível em: <<http://chroniclingamerica.loc.gov>>. Acesso em 10 fev. 2016.

França


Além da imagem da página (Figura 60), o resultado traz o título e o período de circulação da publicação, os trechos em que aparece a palavra-chave inserida na busca. O usuário pode ainda ver informações detalhadas (título, editor, data, origem etc), caso deseje, assim como todos os números daquele jornal ou todos os números resultantes da busca. Os resultados podem ser filtrados, em um menu à esquerda, por época, idioma, autor, assunto etc.


Figura 60 – Detalhe da apresentação de resultados na Gallica



1

Apostolat positiviste : légation positiviste occidentale auprès de la très-sainte ville de Paris - 1903-19??



i Informations détaillées 


Extrait 1 : Parmi les témoignages de sa fidélité, qui lui fit resserrer de plus en plus les liens entre l'Église de Liverpool et celle du Brésil, nous nous bornerons à rappeler son concours décisif, soit au rachat de la Maison de Clotilde, soit à la consécration de celle-ci au culte de l'Humanité, comme vous savez

Extrait 2 : An Brésil, nous avons perdu: en 49/115 (1903), notre jeune coreligionnaire Francisco Bueno Horta Barbosa

Extrait 3 : re vMaitre, dans la politique républicaine,- au Brésil

[Voir tous les numéros](#)

[Voir les numéros correspondant à la recherche](#)



Fonte: Biblioteca Nacional da França. Disponível em: <<http://www.bnf.fr>>. Acesso em: 10 fev. 2016.

Liechtenstein

Os resultados são apresentados em uma lista de *links*, composta pelo nome do jornal, a data completa e a página (Figura 61). Ao lado, o usuário vê o ícone da página gravada em PDF. Em um menu à esquerda, pode-se refinar a busca por período, título e assunto.

Figura 61 – Detalhe da apresentação de resultados na e-Lichtensteinesia

1. [Liechtensteinische Landeszeitung vom 22.12.1866, Seite 4](#) 
2. [Liechtensteinische Landeszeitung vom 5.1.1867, Seite 3](#) 
3. [Liechtensteinische Wochenzeitung vom 7.2.1873, Seite 4](#) 
4. [Liechtensteinische Wochenzeitung vom 21.2.1873, Seite 4](#) 
5. [Liechtensteinische Wochenzeitung vom 10.10.1873, Seite 4](#) 
6. [Liechtensteinische Wochenzeitung vom 20.3.1874, Seite 1](#) 
7. [Liechtensteinische Wochenzeitung vom 12.2.1875, Seite 3](#) 
8. [Liechtensteinische Wochenzeitung vom 2.4.1875, Seite 2](#) 
9. [Liechtensteinische Wochenzeitung vom 3.9.1875, Seite 4](#) 

Fonte: Biblioteca Nacional de Liechtenstein. Disponível em: <<http://www.landesbibliothek.li>>. Acesso em 10 fev. 2016.

México

Os resultados são apresentados de duas maneiras, ou em forma de ficha (imagem, Figura 62) ou em uma lista. Em ambas, o usuário tem acesso a título, data, país, estado, cidade, periodicidade, idioma, página e a informação sobre o acesso (se o PDF está disponível on-line ou somente na Biblioteca Nacional do México). No menu à esquerda, o usuário tem acesso a filtros para refinar os resultados por país, estado, cidade, tipo de acesso, idioma, periodicidade, título e data.

Figura 62 – Detalhe da apresentação de resultados na hemeroteca mexicana



Fonte: Biblioteca Nacional do México. Disponível em: <<http://bnm.unam.mx>>. Acesso em: 10 fev. 2016.

Nova Zelândia

A lista de resultados (Figura 63) é composta por *links*, que direcionam à imagem da página desejada, nome e data da publicação. Não é possível filtrar os resultados.

Figura 63 – Detalhe da apresentação de resultados no Paper Past



Fonte: National Library of New Zealand. Disponível em: <<http://natlib.govt.nz>>. Acesso em: 10 fev. 2016.

Reino Unido

Os resultados são representados pela imagem da página recuperada, o trecho em que aparece a palavra-chave inserida no mecanismo de busca, a data completa e o dia da semana, o título, o condado, país, a especificação do tipo (artigo), o número de palavras e, quando há, as *tags* inseridas pelos usuários cadastrados. É possível filtrar os resultados em um menu localizado à esquerda na página por título, data, região, condado, cidade, tipo e pelas *tags*.

Figura 64 – Detalhe da apresentação de resultados no BNA



Fonte: British Library. Disponível em: <<http://www.bl.uk>>. Acesso em: 10 fev. 2016.

Cada hemeroteca oferece soluções diferentes para a filtragem dos resultados (Quadro 10).

Quadro 10 – Recursos usados na filtragem dos resultados

Hemerotecas	Opções de filtros
Áustria	por tipo de veículo (jornal ou revista), local de publicação, idioma, período de cobertura ou assunto
Austrália	por estado (local), título, categoria, se é ou não ilustrado, década e número de palavras
Bélgica	possível escolher entre visualizar somente a coleção <i>on-line</i> , a coleção física ou as duas
Espanha	<i>não é possível filtrar as informações recuperadas</i>
Estados Unidos	<i>não é possível filtrar as informações recuperadas</i>
França	por época, idioma, autor, assunto
Liechtenstein	por período, título e assunto
México	por país, estado, cidade, tipo de acesso, idioma, periodicidade, título ou data
Nova Zelândia	<i>não é possível filtrar os resultados</i>
Reino Unido	por título, data, região, condado, cidade, tipo e pelas <i>tags</i>

As hemerotecas da Austrália, da Áustria, da Bélgica, dos Estados Unidos e da França oferecem o endereço eletrônico da página selecionada, o que facilita o retorno direto àquele resultado específico.

Nas hemerotecas das BN da Austrália, Áustria, Espanha e Estados Unidos, o usuário pode compartilhar a página recuperada em redes sociais como *Facebook*, *Twitter* e *Google+* ou por e-mail (Quadro 11).

Quadro11 – Compartilhamento de resultados via “Web”

Hemeroteca	Recursos
Austrália	<i>Facebook, Twitter, Digg, Del.icio.us</i>
Áustria	<i>Facebook, Twitter, Google+, E-mail</i>
Espanha	<i>Facebook, Twitter</i>
Estados Unidos	<i>Facebook, Google+, Pinterest, Tumblr, Twitter, Reddit</i>

A possibilidade de copiar o texto do jornal, permitida pelo uso de OCR, é oferecida aos usuários pelas BN da Austrália, Espanha, Nova Zelândia, Reino Unido e Estados Unidos.

7 BRASIL: DO PRIMEIRO JORNAL À HEMEROTECA DIGITAL

A história da imprensa no Brasil começa bem mais tarde do que nos países vizinhos, colonizados por espanhóis¹⁸. Apesar de a primeira prensa ter chegado à América Colonial em 1639 e de algumas tentativas de publicação terem sido empreendidas ainda no século XVII, até a chegada da família real portuguesa ao Brasil, em 1808, as letras impressas eram expressamente proibidas – “pelo poder civil (Ordinário e Desembargo do Paço) e pelo poder eclesial (Santo Ofício)”, segundo Morel (2008, p. 22).

Inicialmente, a finalidade da imprensa era imprimir atos do governo e divulgar notícias interessantes à Coroa, como descreve Lustosa (2004). A instalação da Impressão Régia no Rio de Janeiro ocorreu apenas em 13 de maio de 1808, como apontado no histórico elaborado pela Associação Nacional de Jornais (ANJ)¹⁹.

No mesmo ano em que a corte portuguesa se instalava no Brasil, desencadeando uma onda de cultura e progresso com a abertura dos portos às nações amigas, começava a circular no país o que pode ser considerado o primeiro jornal brasileiro. O *Correio Braziliense* ou *Armazém Literário* (Figura 65), porém, era editado em Londres por Hipólito José da Costa Pereira Furtado de Mendonça (1774-1823), “um brasileiro que nascera no Uruguai, se formara em Portugal, conhecera os Estados Unidos e (...) viveria a maior parte da vida na Inglaterra”, descreve Lustosa (2004, p. 9).

Há muita discussão sobre o *Correio Braziliense* poder ser ou não considerado um jornal brasileiro. Sodré (1999, p. 20) argumenta que o questionamento ocorre não pelo fato de o jornal ser produzido em Londres, mas por tratar sobre os problemas do Brasil mais segundo condições internacionais do que nacionais.

¹⁸ Nelson Werneck Sodré, em *História da imprensa no Brasil*, explica por que a imprensa e as universidades chegaram mais rápido às colônias espanholas. O autor afirma que os portugueses encontraram no litoral americano do Atlântico comunidades primitivas, na fase cultural da pedra lascada. Aquelas que não serviram para ajudar no trabalho foram exterminadas. Nas que aproveitaram o trabalho indígena, preponderou a destruição cultural por meio da catequese jesuítica. A ideia não era exatamente alfabetizar as crianças indígenas, mas destruir nelas a cultura de seus pais. Já na área de colonização espanhola, os europeus encontraram culturas avançadas, já em nível adiantado de complexidade. Ao criar universidades, instituir a imprensa, os colonizadores queriam substituir a cultura encontrada pela sua. Assim, não haveria riscos para o domínio. (SODRÉ, 1999, p. 10-11).

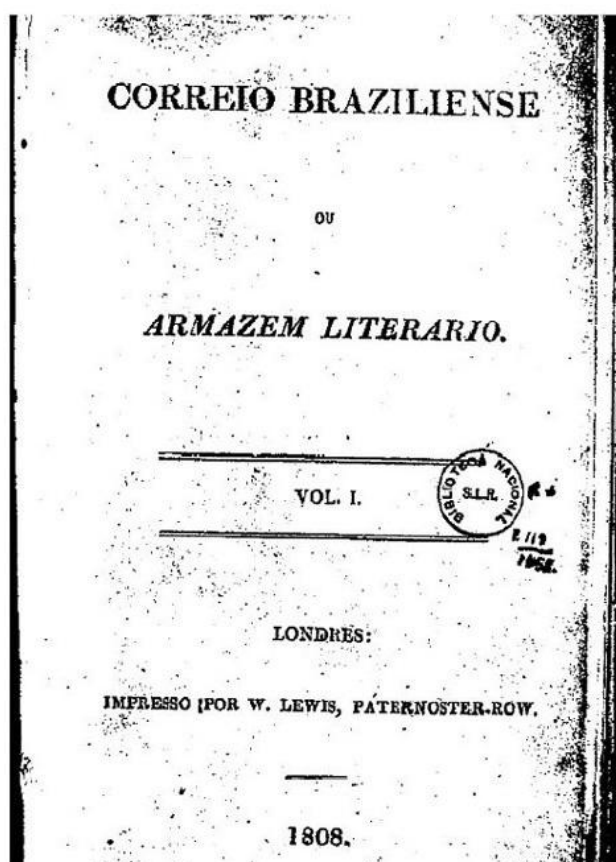
¹⁹ Associação Nacional de Jornais. *História da Imprensa brasileira: dois séculos de história*. Disponível em: <<http://www.anj.org.br/imprensa-brasileira-dois-seculos-de-historia-2>>. Acesso em: 23 out. 2015.

Lustosa, no entanto, argumenta que:

Mesmo cobrindo acontecimentos internacionais, todas as notícias que se imprimiam no *Correio* tinham como alvo o Brasil e seu possível público leitor. Era para informar os brasileiros do que se passava no mundo, para influir sobre seus espíritos, direcionando-os no sentido das ideias liberais, para chamar a atenção para o caráter daninho do absolutismo ou de qualquer forma de despotismo que Hipólito escrevia. (LUSTOSA. 2004, p. 17).

O *Correio Braziliense* circulou de 1º de junho de 1808 a dezembro de 1822, período transcorrido entre a chegada da família real portuguesa, que vinha fugida da onda napoleônica que tomava a Europa, e a proclamação da independência do Brasil.

Figura 65 – Primeira página do *Correio Braziliense*



Fonte: Hemeroteca Digital da Biblioteca Nacional do Brasil (2016).

O mensário tinha o tamanho e o formato de um livro, chegando a mais de 100 páginas divididas nas editorias sobre política, comércio e artes, literatura e ciência, além de uma seção chamada miscelânea. “Composto de longos e densos artigos, nos quais a informação era veiculada de forma circunstanciada e analítica em textos que, às vezes, se prolongavam por vários números seguidos”, como descreve Lustosa (2004, p.15).

Na seção *Reflexões sobre as novidades do mês*, Hipólito José da Costa apresentava suas ideias e seus projetos para o Brasil – acredita-se que o jornal era todo escrito pelo jornalista, mas nem todos os artigos eram assinados. Os planos de Hipólito José da Costa previam o progresso do país que ainda se encontrava muito longe de nações desenvolvidas, como os Estados Unidos e a Inglaterra.

Sodré procura recuperar em *História da Imprensa no Brasil* a reação do governo diante da circulação do *Correio Braziliense*. Quando se constatou que o jornal era lido “por todos” no Brasil, assim como por portugueses em Portugal e em outros países europeus, publicou-se no Brasil, em 27 de março de 1809, uma ordem de “apreensão de material impresso no exterior contendo críticas ao governo” (SODRÉ, 1999, p. 25).

A perseguição ao jornal sofreu altos e baixos ao longo dos anos, não tendo o mesmo peso em todos os lugares do Brasil. O interesse pelo jornal também variava de um local para outro. Deve-se considerar que, na época, os letrados não chegavam a 10% da população.

O atraso da imprensa no Brasil, aliás, em última análise, tinha apenas uma explicação: ausência de capitalismo, ausência de burguesia. Só nos países em que o capitalismo se estabeleceu, a imprensa se desenvolveu. A influência do *Correio Braziliense*, pois, foi muito relativa. Nada teve de extraordinário. Quando as circunstâncias exigiram, apareceu aqui a imprensa adequada. E por isso é que só por exagero se pode enquadrar o *Correio Braziliense* no conjunto da imprensa brasileira. Quando começou a circular, com a clandestinidade obrigada ou não a que se submeteu – clandestinidade porque proibido ou clandestinidade porque pouco lido – não se haviam gerado aqui ainda as condições para o aparecimento da imprensa. O que existia era arremedo. Quando surgiram aquelas condições, o *Correio Braziliense* perdeu a razão de ser (SODRÉ, 1999, p. 28).

O *Correio Braziliense* é considerado o primeiro jornal brasileiro, mas a *Gazeta do Rio de Janeiro* foi o primeiro a ser impresso no Brasil. Redigida pelo Frei Tibúrcio José da Rocha, a *Gazeta* “era uma espécie de folha oficial, onde se publicavam os decretos e os fatos relacionados à família real”, como define Lustosa (2004, p. 20). Após a morte de Rodrigo de Souza Coutinho, responsável direito pela *Gazeta*, a redação foi assumida por Manuel F. de Araújo Guimarães, que trabalhava na produção de textos para *O Patriota* e *O Espelho*, como afirma Morel (2008, p. 30). Havia notícias internacionais, mas nada que fosse contra as ideias da Coroa. Existia uma rigorosa censura imposta pela Imprensa Régia.

Diferente do *Correio Braziliense*, com suas mais de 100 páginas, capa dura e preço muito alto, a *Gazeta do Rio de Janeiro* “era embrião de jornal, com a

periodicidade curta, intenção informativa mais do que doutrinária e formato peculiar aos órgãos impressos do tempo, com poucas folhas e preço baixo”, explica Sodré (1999, p. 22). Os dois jornais, porém, partilhavam do mesmo universo de referências comuns, como afirma Morel.

Tanto a *Gazeta* quanto o *Correio* defendiam idêntica forma de governo (monárquica), a mesma dinastia (Bragança), apoiavam o projeto de união luso-brasileira e comungavam o repúdio às ideias de revolução e de ruptura, padronizados pela crítica comum à Revolução Francesa e sua memória histórica durante a Restauração (MOREL, 2008, p. 30).

Nas duas décadas seguintes à vinda da família real portuguesa, poucos novos jornais foram publicados no Brasil. Porém, podem-se encontrar obras importantes, como as publicadas pelos “inconfidentes” Cláudio Manuel da Costa e Tomás Antonio Gonzaga, além de relatos de viajantes luso-brasileiros e portugueses radicados em outros continentes, segundo revisão de Morel (2008, p.32).

Na Europa, no entanto, surgiram diversas publicações criadas com o apoio do governo a fim de combater a influência do *Correio Braziliense*, como *O investigador português*, inicialmente escrito pelos médicos portugueses Vicente Pedro Nolasco da Cunha e Bernardo José de Abrantes e Castro, e *O contemporâneo*, redigido por Manuel Inácio Martins Pamplona.

De acordo com a ANJ, apesar das transformações econômicas, sociais e políticas ocorridas no Brasil desde a chegada da família real, a situação da imprensa não se alterou consideravelmente antes de 1821, ano em que D. João foi obrigado a retornar a Portugal. Jornais independentes começaram a circular em maior número no Rio de Janeiro em 1821. Uma dessas publicações foi o *Conciliador do Reino Unido*, curiosamente publicado pelo diretor da censura da Imprensa Régia, o baiano José da Silva Lisboa, mais tarde denominado Visconde de Cairu.

Outros jornais criados foram *O Amigo do Rei e da Nação*, *O Bem da Ordem*, *Idade d’Ouro do Brasil*, *Semanário Cívico* e o *Jornal de Anúncios*. Em defesa da liberdade de imprensa, Joaquim Gonçalves Ledo e Januário da Cunha Barbosa criaram o *Revérbero Constitucional Fluminense*, primeira publicação que não passava pelo crivo de censores. O jornal tinha clara pretensão de incentivar a independência do Brasil. Quando o Brasil se libertou de Portugal, o *Revérbero* parou de circular.

A época era de grande tensão, resultando na libertação do Brasil em relação a Portugal, como comenta Lustosa em *O nascimento da imprensa brasileira*. Com a volta da família real a Portugal, surgia uma ameaça de regresso a uma situação que não

poderia ser suportada pela classe dominante brasileira, como afirma Sodré (1999, p. 45), o Brasil novamente subjugado por Portugal.

A conquista anterior [a independência], que parecera simples dádiva, fica em perigo, e o perigo une. Para unir, é preciso mobilizar. Para mobilizar, é preciso despertar a opinião. Para despertar a opinião, é preciso imprensa. Ela tem, então, a sua primeira fase autêntica, entre nós, quando os episódios vividos entre o retorno da corte joanina a Portugal e a mudança de atitude do príncipe regente D. Pedro marca a evolução dos acontecimentos (SODRÉ, 1999, p. 45).

O grito de independência começou a ser dado com a publicação do folheto chamado *Despertador Braziliense*, que ia contra as exigências de Portugal, entre elas o retorno imediato do príncipe regente D. Pedro e o restabelecimento do monopólio comercial com Portugal.

O Brasil de 1821 nada mais tinha a ver com aquela colônia de 1808. O *Despertador Braziliense*, junto com o *Revérbero* e o *Malagueta*, davam início ao movimento do Fico, uma luta pela permanência de D. Pedro no Brasil e contra as medidas das cortes, que claramente visavam ao retrocesso, como observa Lustosa (2004, p. 25). Em 9 de janeiro de 1922, D. Pedro declarava que “para o bem do povo e felicidade geral da nação” ele permaneceria no Brasil.

Figura ilustre na história do jornalismo brasileiro, Cipriano Barata iniciou em 1822 a série de *Sentinelas da Liberdade*. Mesmo preso, de onde estivesse, ele continuava editando seu jornal. Barata, “pioneiro da imprensa libertária no Brasil”, como afirma Sodré (1999, p. 67), tinha claras tendências republicanas, fundando o primeiro jornal sobre o tema no país. Os jornais tinham vida curta nesse período, chegando a apenas poucos exemplares. O *Sentinela*, porém, inspirou muitos outros.

Especula-se que o próprio D. Pedro I costumava escrever contundentes artigos, mas sempre sob pseudônimo, como os publicados em *O Espelho*, jornal criado em 1821 pelo coronel Manuel Ferreira de Araújo Guimarães, ex-editor da *Gazeta do Rio de Janeiro* de 1811 a 1821. Acredita-se que textos virulentos contra opositores foram escritos pelo imperador. Um desses textos teria incitado agressões a Luiz Augusto May, editor do *Malagueta*, notório inimigo de D. Pedro.

No mesmo mês da proclamação da independência, surgiria o *Correio do Rio de Janeiro*, criado com o objetivo claro de encabeçar uma campanha pela Constituinte brasileira. Nessa época, como escreve Lustosa (2004, p. 31), não havia legislação para coibir abusos e eram comuns os ataques àqueles que não concordavam com as ideias veiculadas nas páginas dos jornais, especialmente nas seções de cartas do leitor.

Em 1821, Dom João havia assinado um decreto suspendendo provisoriamente a censura, mas a liberdade de imprensa seria estabelecida somente em 1824 na constituição assinada por D. Pedro. Antes disso, o ministro do reino José Bonifácio de Andrada e Silva, aproveitando o poder que lhe fora concedido, fechou os jornais que apresentavam ideias contrárias às suas – tal devassa aos jornais passou à história com o nome de *Bonifácia*.

A imprensa dividiu-se em dois grupos, prós e contra José Bonifácio, pelo menos até julho de 1823, quando D. Pedro destituiria o seu ministro – José Bonifácio ainda teria um papel importante no país, ao se tornar tutor de D. Pedro II, futuro príncipe regente que assumiria a função com cinco anos de idade.

De acordo com o levantamento realizado por Lustosa, nas décadas seguintes surgiram inúmeras publicações. Em Minas Gerais, em 13 de outubro de 1823 surgiu o primeiro jornal: *O Compilador Mineiro*, publicado em Ouro Preto. Os primeiros jornais impressos em São Paulo foram *O Farol Paulistano*, em 7 de fevereiro de 1827, e *O Observador Constitucional*, de 23 de outubro de 1829, este assinado por Libero Badaró. Anterior, manuscrito, havia o *Mestrinho*, de Antônio Mariano de Azevedo.

Nos anos 1830, como relata Sodré (1999, p. 162), surgiram os pasquins, folhas volantes anônimas que continham apenas um artigo por edição. Todos os pasquins traziam uma epígrafe. “Não havia pequeno jornal que fugisse ao gosto de estampar, em prosa ou verso, um motivo qualquer com relação ao programa ou princípio ou propósito a que obedecia.” (SODRÉ, 1999, p. 162) Como destaca o autor, o papel do pasquim na história da imprensa brasileira foi de inequívoca e fundamental importância.

Nessa época, D. Pedro II foi empossado aos 14 anos. O reinado durou 49 anos.

Da imprensa pioneira, de acordo com a ANJ, seguem em circulação o *Diário de Pernambuco*, de 7 de novembro de 1825, o mais antigo do Brasil atualmente, o *Jornal do Comércio*, fundado no Rio de Janeiro em 1º de outubro de 1827, e o *Monitor Campista*, de Campos dos Goytacazes (RJ), criado em 4 de janeiro de 1834.

A história do *Jornal do Comércio*, segundo Sodré, se confunde com a do reinado. Tida como não partidária, a publicação tem ação intensa e eficaz nos terrenos político, literário e artístico. O tom do jornal é conservador. Contemporâneo do *Jornal do Comércio* foi o *Correio Mercantil*, mas este adotou uma posição política. O *Correio* pertencia a Joaquim Francisco Alves Branco Muniz Barreto, mas a direção do jornal estava a cargo de seu genro, Francisco Otaviano de Almeida Rosa.

O diretor reuniu em seu jornal grandes intelectuais brasileiros, como Manuel Antônio de Almeida (que publicaria no *Correio Mercantil*, entre junho de 1852 e 31 de julho de 1853, um romance de costumes: *Memórias de um sargento de milícias*). José de Alencar também foi um dos chamados por Francisco Otaviano, antigo colega da faculdade de direito. Inicialmente, como relata Sodré (1999, p. 190), Alencar assinava com propriedade a sessão forense, mas aos poucos começou a escrever crônicas sobre os acontecimentos da semana.

O Brasil vivia um período de novos negócios. A lei que extinguiu o tráfico negreiro levou a “criação de bancos emissores, sociedades em comanditas, sociedades colonizadoras, empresas de estradas de ferro, e como complemento natural, a especulação, a agiotagem, o falso luxo”, como cita Sodré (1999, p. 190-191) em sua revisão da história da imprensa. O autor escreve também que os textos de José de Alencar refletiam as mudanças pelas quais passava o país. O autor noticiava, entre outros assuntos, a inauguração das corridas de cavalo, os partidos que se formavam após as récitas do teatro lírico e o aparecimento das máquinas de costuras, às quais era contra porque acabariam com o trabalho caseiro.

O escritor saíria do *Correio Mercantil* em 1855, mas suas crônicas teriam continuidade no *Diário do Rio de Janeiro*, do qual se tornou editor-chefe e no qual publicaria alguns de seus romances mais famosos, como *O guarani* e *A viúvinha*. Assim como José de Alencar, outros homens das letras dedicaram-se à imprensa – Gonçalves Dias, Saldanha Marinho, D. J. Gonçalves de Magalhães, Joaquim Serra, Joaquim Manuel de Macedo, Machado de Assis.

Se nos anos 1850, o poder imperial estava no auge, nos 1860 iniciavam-se mudanças:

A crise de 1857 repete-se em 1864, em circunstâncias ainda mais profundas; surge o conflito militar palatino, de sérias consequências na vida política do país; a suspensão do tráfico negreiro, dez anos antes, mostra claramente os seus efeitos; a quietude é ameaçada, as tormentas vão se acumulando. Ao fim da década de sessenta, com a guerra [do Paraguai] terminada, tudo indica o início de uma fase nova, com reformas que se impõem e não podem ser proteladas; a luta política se acirra; a imprensa retoma o fio de sua história, interrompido pela Maioridade [de D. Pedro II]. Vai começar a agitação (SODRÉ, 1999, p. 201).

Agitação significava peças de teatro em profusão, lançamentos de livros, a moda ganhando espaço e o clamor por uma nova ordem política. Havia no ar uma insatisfação com o reinado de D. Pedro II e a vontade da adoção da República. Esse cenário se refletia nos jornais da época, muitos deles escritos por personalidades que

entrariam para a história, como Olavo Bilac, Raul Pompéia, Quintino Bocaiúva e Aristides Lobo. Sodré (1999, p. 212) conta terem sido criados mais de 20 jornais republicanos entre 1870 e 1872.

Como reporta a ANJ, o Brasil nesta época era essencialmente rural, área em que morava 90% da população, formada por 85% de analfabetos. Desta forma, a imprensa ficava restrita aos centros urbanos. A postura tolerante de D. Pedro II colaborava para a expansão da imprensa nas capitais.

No final do século XIX, como aponta a ANJ, os jornais eram pautados por quatro temas especialmente: monarquia x república e abolicionistas x partidários da ordem escravocrata. O país entra na era das ferrovias, assim como passa a usar o telégrafo – o primeiro telégrafo elétrico foi instalado no Brasil em 1852, ligando o Palácio Real da Quinta da Boa Vista e o Quartel General do Exército. Em 1874, foi inaugurado o cabo submarino que ligava Recife a Londres, o que permitia aos jornais o recebimento de informações sobre os principais acontecimentos registrados no mundo.

Em 1874, a Reuters-Havas instalou-se no Rio e o *Jornal do Comércio* publicou em 1877 os primeiros telegramas distribuídos pela agência. Logo todos os jornais adotaram o novo recurso. Nesse mesmo período, ganharam destaque e espaço as revistas ilustradas, como *O Arlequim*, *Vida Fluminense*, *A Semana Ilustrada* e *O Mequetrefe*. Uma das mais notórias foi a *Revista Ilustrada*, que começou a circular em 1º de janeiro de 1876.

Sua popularidade foi logo muito grande; aparecia aos sábados, vendida a 500 réis o exemplar, com assinaturas anuais a 12\$000 e 20\$000, para a capital e para o interior, respectivamente. As assinaturas podiam ser tomadas na rua da Assembleia, 44, onde estava a sua oficina litográfica a vapor, ou à rua do Ouvidor, 65, na Livraria Garnier. A tiragem atingiu 4000 exemplares, índice até aí não alcançado por qualquer periódico ilustrado da América Latina, regularmente distribuído em todas as províncias e nas principais cidades do interior, com assinatura por toda parte (SODRÉ, 1999, p. 217).

É declarada a abolição da escravatura em 1888. O império chega ao fim com a proclamação da república em 15 de novembro de 1889. D. Pedro II e a família são enviados à Europa em exílio. O início da história do Brasil como uma república é marcado pela caça aos jornalistas e jornais que ainda defendiam a monarquia e que fizessem qualquer comentário desabonador sobre Deodoro da Fonseca ou Floriano Peixoto.

Segundo Sodré (1999, p. 275), a passagem do século assinala no Brasil a transição da pequena à grande imprensa. Os pequenos jornais, de estrutura simples,

cedem lugar às empresas jornalísticas com infraestrutura mais profissional, parque gráfico equipado com o necessário para rodar jornais cada vez com formato mais profissional. As relações com os leitores, com os assinantes e, especialmente, com os anunciantes, são alteradas. “O jornal será, daí por diante, empresa capitalista, de maior ou menor porte”, observa Sodré (1999, p. 275). O tempo do jornal como uma aventura individual, para defender um ponto de vista, é passado.

Nesse período de transformações, a imprensa conheceu múltiplos processos de inovação tecnológica que permitiram o uso de ilustração diversificada – charge, caricatura, fotografia –, assim como aumento de tiragens, melhor qualidade de impressão, menor custo de impresso, propiciando o ensaio da comunidade de massa (ELEUTÉRIO, 2008, p. 82).

A imprensa fervilha no Rio de Janeiro e em São Paulo no início do século XX. O centro do Rio de Janeiro passa por enormes reformas arquitetônicas. Em 7 de abril de 1908, Gustavo de Lacerda funda a Associação Brasileira de Imprensa (ABI). Surgem publicações voltadas a comunidades de imigrantes, os quais chegaram em grande número desde o final do século XIX. Somente em São Paulo, havia mais de 30 jornais em sete idiomas estrangeiros.

Nos primeiros anos do século XX surgiram revistas de cunho humorístico, muitas delas com duração efêmera. Destacam-se *O Malho* (1902), *O Tico-Tico* (1905) e *Fon-Fon* (1907), *O Filhote da Careta* (1910) e *Kosmos* (1904). Editada por cinco anos, a *Kosmos* é tida como “paradigma do mais bem acabado empreendimento entre as revistas consideradas modernas”, segundo pesquisa de Eleutério (2008, p. 90). A revista representava a visão de progresso daqueles tempos eufóricos, uma espécie de *Belle Époque* brasileira.

Dirigida por Mario Behring, em cada edição trazia a colaboração de renomados homens das letras da época, a exemplo do crítico João Ribeiro, o poeta Olavo Bilac, o romancista Coelho Neto, o teatrólogo Artur Azevedo, o jornalista João do Rio (Paulo Barreto), o historiador Capistrano de Abreu, o crítico de arte Gonzaga Duque, entre outros (ELEUTÉRIO, 2008, p. 90).

Tais publicações traziam fotografias da cidade em transformação e dos avanços pelos quais passavam não apenas as grandes cidades, mas seus habitantes. Havia fotografias mostrando as transformações e ilustradores certeiros em suas caricaturas e charges. Nesse cenário inovador, surgiram também revistas para crianças, como a *Tico-Tico* (1905-1962), “primeira no Brasil a publicar histórias em quadrinhos em série e toda a sorte de entretenimento para a garotada, com adivinhações, pequenos jogos, passatempos e quebra-cabeças”, como descreve Eleutério (2008, p. 92).

Novamente a situação política do país mostrava-se tensa. De um lado Rui Barbosa e sua campanha civilista, de outra o marechal Hermes da Fonseca, que seria eleito em 1910. Nesse período surgiria *A Noite*, de Irineu Marinho. Fortes na capital e com posições políticas firmes, existiam o *Correio da Manhã*, com Edmundo Bittencourt, e *O País*, dirigido por João Lage. Na profunda revisão realizada por Sodré, é no início do novo século que são citadas jornalistas mulheres, como Carmem Dolores, pseudônimo de Emília Moncorvo Bandeira de Melo, redatora de *O País*.

Seguindo uma prática iniciada décadas antes, os jornais seguiam publicando em capítulos as obras de seus jornalistas, que mais tarde se tornariam importantes escritores brasileiros. *O triste fim de Policarpo Quaresma*, de Lima Barreto, foi publicado no *Jornal do Comércio* a partir de 11 de agosto de 1911. Novidade na época foi o lançamento de uma página inteira de esportes pelo *Jornal do Brasil* em 1912.

Em 1915, chegou ao Rio, vindo de Pernambuco, Assis Chateaubriand para assumir o cargo de editor-chefe do *Jornal do Brasil*. Este e o *Jornal do Comércio* eram as grandes empresas da época, mas “o órgão popular por excelência continua a ser o *Correio da Manhã*, folha de oposição, vibrante, escandalosa às vezes, veemente sempre”, aponta Sodré (1999, p. 347).

O autor destaca duas figuras dominantes na imprensa da época. Alcindo Guanabara, mais jornalista do que escritor, capaz de escrever com clareza e brilho, mas com pouco discernimento para escolher por quais causas lutar, sendo um retrato da maneira de se posicionar dos jornalistas daquele tempo. Paulo Barreto, mais escritor do que jornalista, “presenciou a transição da folha quase puramente literária para a folha em que a informação começava a ganhar destaque, acabando por dominar a opinião” (SODRÉ, 1999, p. 352). Conhecido como João do Rio, Paulo Barreto retratava nas páginas dos jornais o perfil do novo carioca.

Em 1923, o rádio chegou ao Brasil, trazido por Edgard Roquette-Pinto. Inicialmente eram feitas transmissões de programas de entretenimento. A veiculação de notícias ainda levaria alguns anos para acontecer. De acordo com Eleutério (2008, p. 92), os novos meios de comunicação – rádio e cinema – propiciaram o surgimento de novas revistas, dedicadas às novas mídias, como *Cena Muda* (1921) e *Cinearte* (1923).

Como descreve a ANJ, nesse meio tempo, os jornais se profissionalizaram, inserindo máquinas de escrever às redações, assim como linotipos para acelerar a composição e rotativas que permitiam aumentar as tiragens e melhorar a impressão.

Inúmeros jornais foram criados nesse período da República Velha (1889-1930), como *O Globo* (Rio de Janeiro/RJ), *A Notícia* (Joinville/SC), *Folha da Noite* (hoje Folha de S. Paulo, São Paulo/SP), *Gazeta do Povo* (Curitiba/PR) e *Estado de Minas* (Belo Horizonte/MG).

Os jornalistas também passaram a ser remunerados pelos artigos. Como escreve Eleutério (2008, p. 94), o pagamento aos articulistas cresceu, mas não o número de leitores pagantes dos jornais. Foi assim que a publicidade começou a desempenhar importante papel na manutenção das publicações.

Os anúncios também atendiam às necessidades dos leitores, ávidos por novidades e dispostos a consumir os produtos trazidos pela industrialização e urbanização. Muitas das propagandas eram escritas por autores renomados, como Olavo Bilac, Emilio de Menezes e Bastos Tigre. A maior circulação de propagandas desencadeou o desenvolvimento das agências de publicidade no país.

Em São Paulo, desenvolviam-se temáticas modernistas, com a criação das revistas *Klaxton* (1922-1923), *Revista de Antropofagia* (1928-1929) “e uma profusão de páginas revisteiras modernistas que se multiplicavam Brasil afora”, como afirma Eleutério (2008, p. 96).

Uma das figuras marcantes desta época foi Oswald de Andrade, que além de colaborar com revistas modernistas, foi redator e crítico teatral do *Diário Popular* e correspondente em Paris para o *Correio da Manhã*.

Cohen (2008, p. 104) destaca o aumento de títulos impressos expostos em “locais de jornais”, como tabacarias, estações ferroviárias, charutarias, livrarias e quiosques. A distribuição das publicações também foi ampliada nas primeiras décadas do século XX. Assinaturas eram vendidas por agentes espalhados Brasil afora.

O desdobramento do setor traduziu-se na diferenciação entre jornais e revistas: ao primeiro, normalmente diário e vespertino, caberia a divulgação da notícia, o retrato instantâneo do momento, abrangendo desde as disputas políticas até o descarrilamento do trem de subúrbio. À revista reservava-se a especificidade de temas, a intenção de aprofundamento e a oferta de lazer tendo em vista os diferentes segmentos sociais: religiosas, esportivas, agrícolas, femininas, infantis, literárias ou acadêmicas, essas publicações atendiam a interesses diversos, não apenas como mercadorias, mas ainda como veículos de divulgação de valores, ideias e interesses (COHEN, 2008, p. 104).

De 1930 a 1945, segundo resumo da ANJ, a imprensa acompanhou a evolução do quadro político brasileiro, que passou por um governo provisório, pela revolução de 1932, pela criação da Constituição de 1934 e pelo estabelecimento do Estado Novo em

1937, quando o exercício da liberdade de imprensa desapareceu.

Com tais ocorrências, o jornalismo político novamente tornou-se tema central da imprensa brasileira. O período, compreendido entre 1945 e 1963, que incluiu o governo desenvolvimentista de Juscelino Kubitscheck, foi marcado pelo aumento da importância da receita publicitária nos meios de comunicação. Assinatura e venda avulsa deixavam de ser a principal fonte de renda dos jornais.

Com o fim da Segunda Guerra, ocorreu um desenvolvimento baseado nas tecnologias do que haviam sido criadas durante o conflito. Além de novos equipamentos, houve também uma mudança na maneira de se escrever e diagramar os artigos: adotou-se o método da pirâmide invertida, com busca por objetividade no texto e leiautes mais atraentes. As redações passaram a se organizar por editorias, passando por um processo de profissionalização.

A imprensa, por decreto, tornou-se um serviço público, sujeita ao controle estatal. O Departamento de Imprensa e Propaganda (DIP) foi criado em 1939, com o objetivo de censurar a produção jornalística, cultural e de entretenimento, produzir conteúdos e controlar o abastecimento de papel.

Todos os impressos periódicos precisaram ser registrados no DIP. Mais de 30% das autorizações foram negadas. O DIP também gerava conteúdo, utilizado por diversos matutinos, semanários e mensários, como afirma Luca (2008, p. 172). Até mesmo a liberação do papel para a impressão dos jornais estava alinhada à defesa do governo. Somente quem era a favor de Getúlio Vargas tinha sua carga de papel isenta de taxas alfandegárias.

A virada dos anos 1950 foi marcada pela criação de dois jornais que polarizariam o mundo da imprensa nas décadas seguintes: *Tribuna da Imprensa* (1949), de Carlos Lacerda, e *Última Hora* (1951), sob o comando de Samuel Wainer. Ambos trabalharam nos jornais de Assis Chateaubriand, e chegaram a morar juntos no início da carreira jornalística. Enquanto a *Tribuna* mantinha-se contra o governo, na *Última Hora* havia sempre textos a favor.

Desde os anos 1950, o *Última Hora* manteve um quadro invejável de colunistas, repórteres e editores: Nelson Rodrigues, Octavio Malta, Adalgisa Nery, Dorian Jorge Freire, Stanislaw Ponte Preta (Sérgio Porto), Ricardo Amaral, Jô Soares, Paulo Francis e Ignácio de Loyola Brandão, segundo Laurenza (2008, p.201). O jornal chegou a ser editado em nove capitais brasileiras.

Diversos outros jornais surgiram nesse período, como o *Correio Braziliense* (Brasília/DF), *Jornal NH* (Novo Hamburgo/RS), *O Dia* (Rio de Janeiro/RJ) e o *Pioneiro* (Caxias do Sul/RS).

Uma das personalidades mais influentes nesse período foi o advogado, jornalista, empresário, fazendeiro, senador, ex-embaixador do Brasil no Reino Unido, Francisco de Assis Chateaubriand Bandeira de Mello. Chatô foi dono de inúmeros jornais, reunidos sob o guarda-chuva dos Diários Associados, emissoras de rádio e TV e da Agência Meridional de Notícias.

Foi criação de Chatô o fenômeno chamado *O Cruzeiro*, revista “com média de 720 mil exemplares semanais (número observado em 25 de agosto de 1954, um dia após o suicídio de Getúlio Vargas), em um país de quase 52 milhões de habitantes, predominantemente rural e semianalfabeto”, segundo Laurenza (2008, p. 181). O prestígio da revista *O Cruzeiro* durou até o início dos anos 1960, quando o posto foi assumido pela revista *Manchete*, de Adolpho Bloch, preferida pelo então presidente Juscelino Kubitschek. “Os Diários Associados sobreviveram bem à ditadura de Getúlio Vargas (1937-1945). Viveram seu ápice no período democrático de 1946-1960, mas sucumbiram ao fechamento político de 1964.” (LAURENZA, 2008, p. 182)

Por obra de Chateaubriand, Getúlio Vargas voltou à presidência do Brasil, em 1950, com 48,7% dos votos. Tudo começou com uma entrevista realizada por Samuel Wainer, então repórter dos Diários Associados. Aos poucos, Chatô foi inserindo o nome de Vargas nas manchetes de seus jornais, não como um líder político, mas alcunhado como o líder das massas.

Eleito Vargas, o jornalista apresenta a conta, tornando-se senador por seu estado, como descreve Laurenza (2008, p. 184) O envolvimento com a política se estenderia até os primeiros anos da década de 1960, quando sofreu um acidente vascular e perdeu a mobilidade. Os artigos que publicava em seus jornais eram escritos por meio da ajuda de uma engenhoca, a qual deixaria de usar no dia 20 de março de 1968. Morreria 15 dias mais tarde.

Quando João Goulart foi deposto e os militares tomaram o poder, a imprensa apoiou o novo cenário. Com o passar dos anos, porém, jornais adotaram posturas críticas em relação ao regime militar. Com grande número de lideranças governamentais e intelectuais em exílio, os jornais perderam importantes contribuições em suas páginas. Samuel Wainer foi um dos exilados, deixando seu jornal aos cuidados de colegas.

Quando retornou ao Brasil em 1968, o país passava por um período negro. O governo acabara de instituir o Ato Institucional nº 5 (AI-5), “suspendendo de vez as garantias institucionais, aumentando as prisões e torturas dos opositores, fechando o Congresso por mais de um ano”, como escreve Laurenza (2008, p. 203). Samuel Wainer morreu em 1980.

A história da imprensa no Brasil também engloba a produção de revistas, especialmente a partir dos anos 1950 – antes disso destacam-se a já mencionada *O Cruzeiro* e a feminina *A Cigarra*, que publicava moldes para costurar. Victor Civita lançou o *Pato Donald* em junho de 1950. “Nascia ali a empresa editorial que dominaria o mercado em poucos anos”, escreveu Corrêa (2008, p. 207) em referência à Editora Abril. A revista *Manchete*, de Adolpho Bloch, começou a ser veiculada em 1952, mesmo ano de criação de *Capricho* por Civita.

Capricho concorria com *Grande Hotel*, ambas trazendo fotonovelas. *Manequim* enfrentava *A Cigarra*. Os personagens Disney pretendiam ocupar o lugar das revistinhas com histórias de *Tom & Jerry*. Os maiores concorrentes Civita eram Assis Chateaubriand, dos Diários Associados, Roberto Marinho, da Rio Gráfica, Adolpho Bloch, da Bloch Editores, e Adolfo Aizem, da Ebal.

Nos anos 1960, surgiram as especializadas *Quatro Rodas* (com artigos e uma lista com os preços dos carros) e *Claudia* (com assuntos dedicados às leitoras: educação dos filhos, relações com o marido, controle de natalidade e problemas com ela mesma). As páginas de *Claudia* traziam textos de Vinicius de Moraes, Fernando Sabino, Paulo Mendes Campos, Marina Colasanti, Rubem Braga e Carmen da Silva, autora da coluna inovadora *A arte de ser mulher*.

Outra publicação de destaque nos anos 1960 foi a revista *Realidade*, de 1966, discutindo temas considerados tabus à época e reunindo também escritores de renome, como Carlos Drummond de Andrade, Adoniran Barbosa, Nelson Rodrigues, Plínio Marcos, Paulo Francis e Carlos Lacerda. A revista durou 10 anos.

Realidade retratou um Brasil que se transformava, tratando de assuntos que, em pleno regime militar, sob censura, não apareciam em outras revistas: a maconha, o clero de esquerda, o casamento de padres, o racismo, a fome. Mostrou um país que também não saía com profundidade na imprensa da época: as mazelas do Nordeste, uma Amazônia desconhecida, os problemas das grandes cidades (CORRÊA, 2008, p. 216).

No contexto das revistas, destaque imediato ganhou a revista *Veja*, publicada desde 1968 pela Editora Abril. Segundo Corrêa (2008, p. 218), com talvez a mais cara

campanha de propaganda jamais feita para uma revista no Brasil, *Veja* esgotou 700 mil exemplares na primeira edição – “e foi caindo, caindo, até chegar abaixo dos 100 mil” nas edições seguintes. A revista cheia de texto inaugurava o gênero das revistas semanais de informação, seguindo o modelo das americanas *Time* e *Newsweek*. Diferente da *Manchete*, colorida e repleta de imagens, a *Veja* não foi bem compreendida pelo público. Foram necessários seis anos para sair do vermelho – mais a criação de um departamento para angariar assinaturas e a persistência de Roberto Civita, o filho de Victor Civita, encarregado pela revista.

Ainda no período de ditadura, uma imprensa alternativa satirizava a situação vivida no Brasil. Exemplos disso foram o *Pif Paf*, editado por Millôr Fernandes em 1964, e *O Pasquim*, que reunia textos sobre política, cultura e humor. A censura, porém, havia retornado com todas as forças, especialmente após a edição do AI-5, no dia 13 de dezembro de 1968, como relata a ANJ.

Somente a partir de 1974, começou uma lenta abertura, que se completaria em 15 de março de 1985. A promulgação de nova Constituição ocorreu em 1988, na qual se consolidou o princípio da liberdade de imprensa. No entanto, somente em 2008, o Supremo Tribunal Federal suspendeu a vigência de dispositivos da antiga Lei de Imprensa.

Os anos 1980 encerraram-se com a primeira eleição direta a presidente da República. Fernando Collor de Mello, do recém-criado Partido da Reconstrução Nacional (PRN) disputou uma eleição polarizada com Luiz Inácio Lula da Silva, do Partido dos Trabalhadores (PT). “A cobertura política foi colocada à prova, como nunca, logo no primeiro ano do governo Fernando Collor de Mello”, escreve Villaméa (2008, p. 260).

Reportagens sobre PC Farias ganharam as páginas das revistas semanais e dos jornais, assim como os desvios de verba realizados pela primeira-dama Rosane Collor. A gota d’água foi a matéria de capa de *Veja: Pedro Collor conta tudo*, “na qual o irmão do presidente afirmava com todas as letras que PC Farias era o testa de ferro de Fernando Collor”, registra Villaméa (2008, p. 262).

Curiosamente, o desfecho da desavença entre Pedro Collor e PC Farias deu-se por meio do envolvimento dos dois com a imprensa. Pedro Collor era diretor das Organizações Arnon de Mello, que mantinham a *Gazeta de Alagoas*, entre outros veículos de comunicação. PC Farias decidiu criar seu próprio jornal, a *Tribuna de*

Alagoas, contrariando Pedro Collor. “Uma disputa empresarial acabou jogando luz no esquema de corrupção instalado em Brasília” (VILLAMÉA, 2008, p. 262).

Com a televisão conquistando cada vez maior espaço e se firmando como o meio de comunicação de massa em nível nacional, os jornais passaram por transformações. Os jornais vespertinos desapareceram, deixando espaço apenas para os matutinos. Várias publicações encerraram suas atividades, mas os grandes viram suas tiragens aumentarem, assim como investiram em tecnologias mais modernas.

A informatização das redações foi iniciada nos anos 1980. O primeiro grande jornal a contar com computadores foi a *Folha de S. Paulo* em dezembro de 1982. Já o *Estado de S. Paulo* foi um dos últimos – abandonando as barulhentas máquinas de escrever somente em 1991. O *Jornal do Brasil*, em 1995, foi o primeiro jornal a ter uma versão na “Web”.

Na virada de milênio o aumento do uso de novas tecnologias, como televisão por assinatura e internet, tornou-se ainda mais marcante. Muitos dos jornais brasileiros criaram versões digitais de seus conteúdos. Apesar dessa nova realidade, novos jornais impressos foram criados nos anos 2000, como *Notícias do Dia* (Florianópolis/SC), *Extra* (Rio de Janeiro/RJ) e *O Sul* (Porto Alegre/RS), entre outros.

Em 2014, de acordo com dados da ANJ, circulam no Brasil diariamente 784 jornais. No total, são 5219 jornais no Brasil. Os cinco maiores são: *Super Notícia/MG* (com uma média de circulação de 318.067); *Folha de S.Paulo* (351.745), *O Globo* (333.860), *Daqui/GO* (203.157) e *Extra* (205.994).

7.1 BIBLIOTECA NACIONAL DO BRASIL

A Biblioteca Nacional do Brasil é considerada pela Organização das Nações Unidas para a Educação, a Ciência e a Cultura (Unesco) uma das dez maiores bibliotecas nacionais do mundo, sendo a maior da América Latina.

O acervo atual estimado em nove milhões de itens, conforme consta do site da instituição²⁰, teve origem na coleção da livraria de D. José, organizada pelo Abade de Santo Adrião de Server, Diogo Barbosa Machado, para substituir a Livraria Real, cuja origem remontava às coleções de D. João I e de seu filho D. Duarte, consumidas pelo incêndio que se seguiu ao terremoto de Lisboa de 1º de novembro de 1755.

²⁰ BIBLIOTECA NACIONAL (Brasil). *Apresentação*. Disponível em: <<https://www.bn.br/sobre-bn/apresentação>>. Acesso em: 12 jan. 2016.

Pouco mais de 50 anos depois, a invasão das tropas de Napoleão Bonaparte a Portugal forçaram a transferência da família real e da corte portuguesa para o Brasil. Assim, a origem da Biblioteca Nacional está intimamente ligada à vinda da família real portuguesa ao Brasil, como descreve Morais (2006, p. 91).

Junto à mudança de D. Maria, de D. João e do príncipe regente, viajou a Real Biblioteca de Portugal. O acervo era composto por 60 mil peças, incluindo livros, manuscritos, mapas, estampas, moedas e medalhas. A família e sua coleção chegaram ao Brasil em março de 1808.

O endereço inicial da biblioteca portuguesa foi uma sala do Hospital do Convento da Ordem Terceira do Carmo, que ficava na Rua Direita, hoje Primeiro de Março, no Centro da cidade do Rio de Janeiro. Como descrito por Morais (2006, p. 91), era uma biblioteca rica e versátil, que continha mais de cem incunábulo, entre eles dois exemplares da Bíblia de Mogúncia de 1462.

Contribuiu para tornar o acervo ainda mais robusto a ordem decretada por D. Maria em 1805 de que uma cópia de todos os papéis impressos nas oficinas tipográficas do reino deveria ser enviada à Real Biblioteca. Anos mais tarde, no Brasil, o governo imperial criaria lei semelhante em 1822, mesmo ano em que a Real Biblioteca passaria a se chamar Biblioteca Imperial e Pública.

Oficialmente estabelecida, a Biblioteca Nacional continuou a ter seu acervo ampliado de maneira significativa, por meio de compras, doações e, principalmente, de ‘propinas’, ou seja, pela entrega obrigatória de um exemplar de todo material impresso nas oficinas tipográficas de Portugal e na Impressão Régia, instalada no Rio de Janeiro (BETTIOL, 2008, p. 3).

Tal prática definida na Lei do Depósito Legal (Lei 10.994), de 14 de dezembro de 2004²¹, que teve como base o Decreto de Contribuição Legal assinado em 1907 pelo então presidente Afonso Pena, dispõe sobre a remessa de obras à Biblioteca Nacional em qualquer suporte físico, destinadas tanto à venda quanto à distribuição gratuita.

O objetivo da lei é assegurar o registro e a guarda da produção intelectual nacional, além de possibilitar o controle, a elaboração e a divulgação da Bibliografia Brasileira corrente, bem como a defesa e a preservação da língua e da cultura nacionais, como ressaltado no site da BN. Com essa prática, a Biblioteca Nacional torna-se a guardiã da memória gráfica brasileira.

²¹ BIBLIOTECA NACIONAL (Brasil). *Depósito Legal*. Disponível em: <<https://www.bn.br/sobre-bn/apresentação>>. Acesso em: 12 jan. 2016.

O decreto que garantiu um endereço à Real Biblioteca foi assinado pelo príncipe regente, D. João VI, no dia 29 de outubro de 1810, tida como a data de fundação da Biblioteca Nacional. A abertura ao público aconteceria, porém, somente em 1814. O acervo estava sob os cuidados do frei Gregório José Viegas e do padre Joaquim Dâmaso.

Em 1821, a Família Real retornou a Portugal, tomando consigo grande parte dos manuscritos do acervo da Biblioteca Real, chamados *Manuscritos da Coroa*. Após a Proclamação da Independência em 1822, a aquisição da Biblioteca Real pelo Brasil por 800 contos de réis foi regulada mediante a Convenção Adicional ao Tratado de Paz e Amizade celebrado entre Brasil e Portugal em 29 de agosto de 1825.

A biblioteca teria ainda outro endereço. Em 1855 foi transferida para a Rua do Passeio – onde atualmente funciona a Escola de Música da Universidade Federal do Rio de Janeiro (UFRJ) –, antes de instalar-se na Cinelândia.

O nome Biblioteca Nacional seria dado definitivamente em 1876, mesmo ano em que começam a ser editados os *Anais da Biblioteca Nacional*²². O objetivo do material produzido até hoje é divulgar documentos preciosos, livros raros e peças curiosas, além de publicar manuscritos e trabalhos bibliográficos.

O primeiro catálogo da Biblioteca Nacional foi inaugurado em 1883 tendo como base o sistema Brunet, utilizado à época na França.

Quando a República foi proclamada, em 1889, o imperador D. Pedro II precisou mudar-se para Portugal. Antes de partir, ele doou um conjunto de aproximadamente 100 mil obras à biblioteca e pediu para que seja chamado de *Coleção D. Thereza Christina Maria*, em homenagem a sua esposa. De acordo com informações institucionais, o acervo reunia livros, publicações seriadas, mapas, partituras, desenhos, estampas, fotografias, litografias e outros documentos impressos e manuscritos.

Na virada do século XX, as bibliotecas cariocas adotaram um catálogo coletivo e passaram a seguir a Classificação Decimal Universal (CDU). Nesse período foi lançado o *Boletim Bibliográfico Brasileiro*, dentro das normas da CDU, que evoluiu para a atual *Bibliografia Brasileira*.

A pedra fundamental do novo prédio da Biblioteca Nacional foi lançada em 1905 na Avenida Central, hoje Avenida Rio Branco, no centro do Rio de Janeiro. Com o

²² BIBLIOTECA NACIONAL (Brasil). *Anais da Biblioteca Nacional*. Disponível em: <<http://bndigital.bn.br/acervo-digital/anais-biblioteca-nacional/402630>>. Acesso em 12 jan. 2016.

novo edifício pronto, o acervo foi transferido entre 1909 e 1910. De acordo com a linha do tempo da Biblioteca Nacional, foram necessárias 1332 viagens para transportar as cerca de 400 mil obras. A inauguração ocorre em 1910, exatamente quando a instituição completou 100 anos. Em 1973, o prédio foi tombado pelo Instituto do Patrimônio Histórico e Artístico Nacional (IPHAN).

Foi dentro das dependências da Biblioteca Nacional que se instituiu o primeiro curso de Biblioteconomia do Brasil e da América Latina. As atividades foram iniciadas em 1915 seguindo o modelo da *École de Chartres*, na França.

Em 1978, foi criado o Plano Nacional de Microfilmagem de Periódicos Brasileiros (PLANO), que visava preservar a produção jornalística do país e supervisionar uma rede nacional de microfilmagem²³. No mesmo ano, a instituição adota o formato CALCO, a Catalogação Legível por Computador. Quatro anos mais tarde, a biblioteca integrou-se ao sistema Bibliodata/Calco da Fundação Getúlio Vargas, desenvolvendo um catálogo automatizado em formato MARC.

Em 1990, a Biblioteca Nacional, com sua biblioteca subordinada, Euclides da Cunha, do Rio de Janeiro, e o Instituto Nacional do Livro, com sua Biblioteca Demonstrativa, de Brasília, passaram a constituir a Fundação Biblioteca Nacional (FBN). A partir de 1998, os catálogos são publicados na “Web”. Em 2014, na última atualização da linha do tempo, noticia-se a aquisição do software *Sophia* de automação bibliográfica. Os catálogos passaram a ser migrados para esta nova plataforma.

O acervo, que se encontra armazenado no prédio da Cinelândia e em um anexo situado na região portuária do Rio de Janeiro, é composto pelas seguintes seções, segundo as informações oficiais do site da Biblioteca Nacional brasileira:

- 1) Obras gerais – mais de um milhão de monografias, teses e folhetos do século XVIII aos dias atuais. Parte da coleção pode ser vista em microfilme.
- 2) Música – maior acervo de música da América Latina, com 220 mil peças, que incluem música erudita e popular, autores nacionais e estrangeiros, músicas de diferentes estilos de época.
- 3) Periódicos – maior coleção de jornais e revistas da América Latina, criada por decreto de 1922, que engloba mais de 58 mil jornais, revistas,

²³ BIBLIOTECA NACIONAL (Brasil). *Plano Nacional de Microfilmagem de Periódicos Brasileiros*. Disponível em: < <http://www.bn.br/explore/planos-preservacao/plano-nacional-microfilmagem-periodicos-brasileiros> >. Acesso em 12 jan. 2016.

boletins técnicos e anuários. Coleção pode ser lida no formato original ou em cópias microfilmadas. Parte está disponível on-line na Hemeroteca Digital Brasileira.

4) Obras raras – livros, folhetos, folhas volantes, periódicos que tenham características ímpares, como o autógrafo de personalidade célebre (D. Pedro II, Jorge Amado, Carlos Drummond de Andrade, por exemplo), a encadernação de luxo ou o fato de pertencer a alguma edição especial.

5) Cartografia – 22 mil mapas, manuscritos e impressos, 2.500 atlas, além de monografias e tratados sobre o tema.

6) Iconografia – obras impressas, estampas raras, desenhos, fotografias, impressos efêmeros e bibliografia especializada. Este acervo contém obras assinadas por Albrecht Dürer, Augusto Malta, Carlos Oswald, Emílio Goeldi, Francis Frith, J. Carlos, Jean Baptiste Debret, Johann Moritz Rugendas, Marc Ferrez, Oscar Niemeyer, Thomas Ender, Volpi, entre outros.

7) Referência – existente desde 1944, totalmente aberta ao público, contém dicionários técnicos e gerais, enciclopédias, bibliografias nacionais e estrangeiras, guias, manuais e atlas.

7.2 BIBLIOTECA DIGITAL DA BIBLIOTECA NACIONAL DO BRASIL

A Biblioteca Nacional Digital, a BNDigital, foi criada em 2006 para integrar as coleções digitalizadas da Biblioteca Nacional. A digitalização do acervo, porém, havia sido iniciada ainda em 2001 no contexto de exposições e de projetos temáticos em parcerias com instituições nacionais e internacionais.

Ao digitalizar seu acervo, como descrito em seu site, a Biblioteca Nacional pretende:

- a) Difundir as coleções da Biblioteca Nacional, diversificando sua esfera social (comunidade ou público-alvo) e enriquecendo as possibilidades de acesso e uso;
- b) Promover a salvaguarda do acervo original da Biblioteca Nacional;
- c) Subsidiar atividades culturais e de extensão, em nível local, nacional e internacional, que envolvam o uso da imagem digital como recurso de difusão, segurança, preservação e salvaguarda.

De acordo com as diretrizes apresentadas no site da BNDigital, a seleção do material a ser digitalizado considera o valor histórico ou memorial do item, a importância e a raridade de obras específicas, assim como a relevância de coleções. Entre os critérios estão a escolha de item que constitua o objeto de missão estatutária da Biblioteca Nacional do Brasil, item em domínio público ou autorizado pelo titular dos direitos intelectuais e morais, item muito procurado pelos usuários e item fragilizado em condições materiais. Em 2015, a prioridade foi a digitalização de obras sobre o Rio de Janeiro.

A BNDigital está internamente constituída por três segmentos: Captura e armazenagem de acervos digitais; Tratamento técnico e publicação de acervos digitais; e Programas e Projetos de digitalização e divulgação. De acordo com o site, conta com uma equipe interdisciplinar composta por bibliotecários, historiadores, arquivistas e digitalizadores.

Sobre os padrões para descrição e representação da informação, o esquema de metadados adotado pela BNDigital é o Dublin Core, no qual se identificam 15 elementos: título, criador, assunto, descrição, publicador, contribuidor, data, tipo, formato, identificador, fonte, idioma, relação, cobertura e direitos.

Conforme a biblioteca, a norma observada para a representação dos pontos de acesso de autoria é a Anglo American Cataloguing Rules (AACR2), que especifica os elementos necessários à descrição e à identificação de publicações.

O sistema de classificação utilizado na BNDigital é a Classificação Decimal de Dewey. O vocabulário controlado adotado para indexação é a Base de Terminologia da FBN, a qual segue a estrutura da lista *Cabeçalhos de Assunto da Biblioteca do Congresso dos Estados Unidos*. Na BNDigital, todos os conteúdos são representados de forma bilíngue: em português e inglês.

Ainda, para interoperar com outros sistemas de bibliotecas digitais a BNDigital aderiu ao protocolo da Iniciativa dos Arquivos Abertos OAI-PMH, mecanismo para transferência de dados entre repositórios digitais.

7.3 HEMEROTECA DIGITAL BRASILEIRA

Biblioteca Nacional do Brasil: <https://www.bn.br/>

Biblioteca Nacional Digital: <http://bndigital.bn.br/>

Hemeroteca Digital Brasileira: <http://bndigital.bn.br/hemeroteca-digital/>

A Hemeroteca Digital Brasileira (HDB) é a versão digital da coleção de periódicos da Fundação Biblioteca Nacional, que engloba jornais, revistas, anuários, boletins, entre outras publicações. A coleção de periódicos da Biblioteca Nacional é a mais antiga e completa desse gênero no país, composta por mais de 58 mil títulos.

Há desde exemplares dos primeiros jornais criados no Brasil – como o *Correio Braziliense* e a *Gazeta do Rio de Janeiro*, ambos fundados em 1808 – até periódicos extintos no século XX (como a revista *Fon-Fon* e os jornais *Diário Carioca*, *Correio da Manhã* e *Opinião*) ou que não circulam mais em versão impressa, como o *Jornal do Brasil*, que autorizou a disponibilização de seu conteúdo na hemeroteca.

Bettencourt (2013, p. 2) afirma que como parte da política de preservação a BN vem, desde meados da década de 1940, microfilmando o acervo. A partir de 1970, projetos procuraram salvaguardar coleções de periódicos pela microfilmagem continuada. Como já escrevia Braga em 1973, “a importância dos jornais para diferentes estudos, em diferentes áreas, pode ser comprovada pela frequência de circulação – e exatamente o grande número de consultas é um dos agentes causadores de problemas de conservação e recuperação”.

Em 1978, foi criado o Plano Nacional de Microfilmagem de Periódicos Brasileiros. Com essa iniciativa, tornou-se possível reunir, resgatar e mapear grande parte da produção de periódicos no país. Atualmente, segundo Bettencourt, nove mil títulos de periódicos encontram-se microfilmados.

Com a HDB, lançada em 2012 com 10 milhões de páginas, a consulta agora é possível a partir de qualquer aparelho conectado à internet, podendo ser realizada por título, período, local de publicação e palavra-chave. A busca por palavras, associada sempre a título, período e local, é possível devido à utilização da tecnologia OCR, que proporciona aos pesquisadores maior alcance na pesquisa textual em periódicos. Outra vantagem do portal é que o usuário pode também imprimir em casa as páginas desejadas ou salvá-las em formato PDF. Cada página é identificada por uma URL.

Bettencourt (2013) ressalta que para a execução do processo de digitalização do acervo foi desenvolvida uma metodologia própria, contemplando as diferentes etapas

da cadeia criada para a implementação do projeto: seleção, captura, descrição e representação, indexação e disponibilização dos arquivos digitais. A autora destaca que o modelo de interoperabilidade adotado foi o mesmo escolhido para a BNDigital, baseado no protocolo OAI-PMH, pois permite a coleta e o intercâmbio de metadados entre repositórios digitais.

Os critérios de seleção, de acordo com Bettencourt, basearam-se nos seguintes princípios, procurando respeitar a legislação dos direitos autorais:

- a) jornais brasileiros, incluindo aqueles publicados fora do território nacional, como é o caso do *Correio Braziliense* (1808);
- b) jornais em domínio público ou aqueles cujos direitos de publicação foram cedidos à BN como é o caso de *Jornal do Brasil*, *Conjuntura Econômica* e publicações do grupo Diários Associados, que inclui a conhecida revista *O Cruzeiro*;
- c) jornais raros e aqueles mais solicitados pelos usuários para consulta local ou para reprodução.

Na etapa seguinte, de captura ou conversão para o digital, adotaram-se dois critérios, um para jornais em preto e branco e outro para coloridos.

A captura dos periódicos em preto e branco ficou a cargo da empresa DocPro e foi realizada na Biblioteca Nacional a partir do microfilme negativo que é o máster em microfilmagem. Foram utilizados escâneres de microfilmes FlexScan de alta produtividade, capazes de digitalizar vinte mil fotogramas por dia. A etapa de captura, executada nesses equipamentos, é quase completamente automatizada, prescindindo da interferência humana (BETTENCOURT, 2013, p. 4).

Grande parte dos materiais coloridos foram capturados no Laboratório de Digitalização da Biblioteca Nacional, onde se utilizaram escâneres planetários de alta produção marca Zeutschel modelo 12000HQ. Tais equipamentos têm características que visam à salvaguarda do documento original.

Bettencourt destaca que após a captura, ocorrem nomeação, separação e encadernação visual, atividades realizadas por profissionais da informação. O trabalho nesta fase é lento e meticuloso. Na etapa de nomeação, visando à preservação e à disponibilização *on-line*, adotou-se o padrão dos arquivos digitais da BNDigital:

Identificador único do título do periódico, precedido pelo prefixo 'per' que é a sigla convencionada pela BN identificadora do tipo de documento – no caso periódico + ano de publicação + número da edição + número sequencial da página. Essas regras valem tanto para os arquivos gerados a partir do microfilme, quanto para os arquivos gerados a partir de documentos originais (BETTENCOURT, 2013, p. 4).

A partir das imagens digitalizadas dá-se início à etapa de reconhecimento dos caracteres do texto por meio da aplicação do software *Abbyy FineReader 11 Professional*. Por último, a encadernação virtual é realizada fascículo a fascículo, em formato PDF, espelhando o documento original. Assim o leitor poderá ter a sensação de estar folheando uma edição antiga.

Além do modo convencional de indexação do conteúdo utilizado na BNDigital, a Hemeroteca Digital Brasileira utiliza o processo de indexação que contempla todo o conteúdo do documento. Para realizá-lo, utiliza-se o “Inteligenciamento DocPro”, um processo que envolve a pesquisa por aproximação visual, característica principal da tecnologia DocPro, na qual não são guardadas as palavras exatas e sim a aproximação visual de cada uma. Desta forma, as falhas que normalmente acontecem em um OCR comum são minimizadas, o que se traduz em uma taxa de acerto muito superior, como afirma Bettencourt (2013, p. 7).

Na página inicial da Hemeroteca Digital (Figura 66), há notícias, destaques e indicações dos títulos mais recentes incorporados ao acervo, assim como as listagens de jornais antigos.

Figura 66 – Página inicial da Hemeroteca Digital Brasileira

Fonte: Biblioteca Nacional Digital. Disponível em: <<http://bndigital.bn.br/hemeroteca-digital>>. Acesso em: 10 fev. 2016.

Quando o usuário realiza uma busca no site da BNDigital, esta pode ser feita por autor, título, editor e datas e palavras-chave. Já no site da Hemeroteca Digital Brasileira, há também a opção de busca por palavra-chave no conteúdo textual de determinado título, período de tempo específico ou local de publicação (Figura 67).

Figura 67 – Busca avançada da Hemeroteca Digital Brasileira

PERIÓDICO PERÍODO LOCAL

Pesquisa por Periódico

1 - Digite ou escolha um periódico

Selecione...

2 - Escolha um período

3 - Digite para pesquisar

Pesquisar

Fonte: Biblioteca Nacional Digital. Disponível em: <<http://bndigital.bn.br/hemeroteca-digital>>. Acesso em: 10 fev. 2016.

Na página inicial da HDB, são oferecidas aos usuários detalhes sobre mais de 100 publicações. Assim, fica-se sabendo o cenário e o contexto em que determinado jornal foi criado.

Durante o período de 14 de dezembro de 2015 a 10 de janeiro de 2016, a Hemeroteca Digital registrou aproximadamente 90 mil *page views*, como mostram as estatísticas apresentadas no site da Biblioteca Nacional. Mais de 90% dos visitantes eram oriundos das Américas. Dentro do Brasil, a maior parte do Rio de Janeiro (26%) e de São Paulo (23%). Em terceiro lugar, Minas Gerais (9%).

8 BUSCA NA HEMEROTECA DIGITAL BRASILEIRA

Para a realização de buscas no acervo da Hemeroteca Digital Brasileira, de forma a analisar o potencial da mesma para subsidiar a atividade de pesquisa científica especializada, foram coletadas informações junto a pesquisadores da área de História da Ciência e da Saúde. A partir de tais informações, elaborou-se a definição e a formulação das necessidades de informação desse tipo de usuário, que serviram de base para as buscas empreendidas nos jornais brasileiros antigos já digitalizados pela Biblioteca Nacional.

As perguntas englobaram informações sobre a formação acadêmica do pesquisador, sua experiência prévia em buscas em uma hemeroteca, sua pesquisa atual e a expectativa do pesquisador na utilização da hemeroteca nacional.

8.1 NECESSIDADES DE INFORMAÇÃO DOS PESQUISADORES

Neste capítulo são apresentadas necessidades de informação elaboradas a partir das informações colhidas no período de 15 a 30 de outubro de 2015 com pesquisadores da área de História das Ciências e da Saúde, pós-graduandos ou egressos do Programa em História das Ciências e da Saúde (PPGHCS) da Casa de Oswaldo Cruz (COC), pertencente à Fundação Oswaldo Cruz (Fiocruz), no Rio de Janeiro, e por um mestrando em História da Universidade Federal de Santa Catarina (UFSC), em Florianópolis.

Pesquisador 1 (PQ1)

O pesquisador planeja escrever a dissertação de mestrado sobre o sociólogo Hiroshi Saito (1919-1983), primeiro acadêmico a investigar a imigração japonesa no Brasil, especialmente entre 1950 e 1980. Saito foi professor da Escola Livre de Sociologia e Política (ELSP), da Pontifícia Universidade Católica de São Paulo (PUC-SP) e da Universidade de São Paulo (USP). A dissertação terá como base as publicações do sociólogo, mas citações em jornais, caso recuperadas, poderão complementar a descrição sobre a atuação de Hiroshi Saito.

Pesquisador 2 (PQ2)

O pesquisador interessa-se por políticas públicas psiquiátricas e políticas públicas em saúde. Para sua tese, ele pretende analisar a assistência psiquiátrica pública do Ministério da Saúde via Serviço Nacional de Doenças Mentais (SNDM), instituição

criada em 1941. O período de foco do doutorando engloba as décadas de 1960 e 1970. O pesquisador trabalhará com periódicos científicos, da área psiquiátrica no Brasil, mas pretende pesquisar nos jornais da época uma reforma administrativa ocorrida em 1970, para ver a repercussão sobre o assunto.

Pesquisador 3 (PQ3)

O pesquisador tem interesse por epidemias e práticas de cura na Amazônia na segunda metade do século 18 e nos primeiros anos do século 19. Especificamente até 1808. A pesquisa engloba essencialmente documentos coloniais, correspondências trocadas entre autoridades da época, entre autoridades dentro do Pará e entre as autoridades do Pará e as de Lisboa. Caso estendesse o período de análise, poderia considerar usar notícias de jornal sobre as epidemias, doenças e práticas de cura. Detectou doenças como sarampo e varíola (+ malária). O público de estudo é composto por trabalhadores da Fortaleza São José do Macapá, que na época pertencia ao Pará. Os trabalhadores eram negros e índios.

Pesquisador 4 (PQ4)

O pesquisador pretende analisar o processo contrahegemônico das práticas de curas populares encontradas nos estudos de folclore brasileiros e portugueses de boa parte do século XX. Busca descrições de práticas populares de curas, descrições de remédios caseiros, das ervas utilizadas, dos modos de fazer, dos rituais envolvendo tais curas, dos curadores e das pessoas envolvidas, das concepções de doença, saúde e sofrimento dessas pessoas.

Pesquisador 5 (PQ5)

O pesquisador tem a necessidade de demonstrar como cirurgiões atuaram quando eram obrigados a manter vínculos com as forças armadas e as câmaras municipais coloniais. Da Hemeroteca Nacional, pretende usar apenas artigos de *O Patriota* e um documento da *Gazeta de Lisboa*. A base da tese é composta por documentos manuscritos.

Pesquisador 6 (PQ6)

A pesquisadora interessa-se por boticas, funcionários do ultramar, intermediários do tráfico a serviço da cura, do final do século XVIII à primeira metade do século XIX no

Brasil/Angola. A pesquisa trata sobre as curas, os profissionais de cura, a circulação de informação médica entre a América Portuguesa, Luanda e Benguela. Como essas informações circulavam nas possessões ultramarinas. Em jornais antigos, poderia pesquisar sobre os negros envolvidos nos processos de curas, anúncios de sangradores negros, atendimento nos navios negreiros, ervas usadas, procedimentos adotados. Poderia procurar pelos nomes das doenças, como varíola (bexiga), febres, cura, disenteria, febre pútrida, febre perniciosa. Por termos como curandeiro e feiticeiro. Autorização a curandeiros.

Pesquisador 7 (PQ7)

A pesquisa tem como foco a atuação das parteiras no Rio de Janeiro no século XIX, entre os anos 1820-1870. Como insumo para sua dissertação, espera encontrar anúncios de parteiras, e, se possível, com maior número de informações. A arte obstétrica no Brasil do século XIX foi incluída nas escolas de medicina e cirurgia, na Bahia e no Rio, em 1808, por edital de Dom João VI. Primeira parteira formada no Rio de Janeiro no ensino oficial de obstetrícia para mulheres foi Maria Josefina Matilde Durocher (1808-93), conhecida como Madame Durocher. Uma Lei de 1832 cria o curso de parteiras.

Pesquisador 8 (PQ8)

O pesquisador investiga manuais de medicina publicados no século XVIII. Pretende realizar uma análise sobre três autores: Luiz Gomes Ferreira, José Antonio Mendes, João Cardoso de Miranda. Existem vários viajantes, mas Ferreira foi o primeiro. Em jornais antigos, poderia pesquisar sobre esses médicos viajantes, que depois escreveram manuais de medicina. Tais erários eram referência para outros médicos. As vindas dos médicos portugueses ocorreram no período colonial.

Pesquisador 9 (PQ9)

O pesquisador desenvolve um projeto de pós-doutoramento sobre as empresas Instituto Behring de Terapêutica Experimental (IBTE), que foram criadas em quatro capitais da América Latina nos anos 1930, como forma de assegurar o mercado de produtos biológicos de uso humano e animal para o conglomerado de indústrias químico farmacêuticas alemã, IG Farben. As necessidades atuais do pesquisador são saber em que medida havia divulgação dos produtos do IBTE no Rio de Janeiro e em São Paulo e

encontrar notícias sobre a prisão dos dirigentes dessa empresa, após o rompimento das relações diplomáticas entre Brasil e Alemanha em 1942.

Pesquisador 10 (PQ10)

O doutorando pesquisa um grupo modernista dos anos 1920 conhecido como “verde-amarelo” (formado por Menotti del Picchia (1892-1988), Plínio Salgado (1895-1988), Guilherme de Almeida (1890-1969) e Cassiano Ricardo (1895-1974) em contraposição ao movimento nacionalista criado por Oswald de Andrade). Tem interesse em acessar as crônicas publicadas pelos integrantes do grupo verde-amarelo entre 1925 e 1929. A maior parte dessas crônicas foi publicada no jornal paulista *Correio Paulistano*, o qual está disponível na Hemeroteca Digital Brasileira.

Pesquisador 11 (PQ11)

O pesquisador pretende encontrar subsídios para desenvolver um projeto de doutorado tendo como temas caridade, construção das políticas públicas brasileiras e da saúde pública. As necessidades do pesquisador concentram-se em encontrar documentos referentes à atuação do Hospital São Francisco de Belo Horizonte, em relação ao atendimento público de saúde prestado aos desempregados durante o governo Vargas até 1988, com a implantação do Sistema Único de Saúde (SUS).

Pesquisador 12 (PQ12)

A área temática do pesquisador engloba a história moderna e contemporânea. É uma continuação da pesquisa de doutorado que tratou sobre a história da psicanálise no Rio de Janeiro, nas primeiras décadas do século XX. O pesquisador já teve uma experiência bem-sucedida na Hemeroteca Digital brasileira, onde buscou pelos nomes dos autores (Juliano Moreira, Henrique Roxo, Julio Porto-Carrero).

Pesquisador 13 (PQ13)

A pesquisadora interessa-se por história regional do Brasil, com foco no estado do Mato Grosso. Investiga o processo de institucionalização da assistência a alienados em Mato Grosso, no período compreendido entre 1931 a 1957. Necessita de informações sobre a assistência a alienados em Mato Grosso e outros estados da federação no período de 1931 a 1957.

8.2 PROCESSOS DE BUSCA E RESULTADOS

As buscas no acervo da Hemeroteca Digital Brasileira da Biblioteca Nacional do Brasil foram realizadas em janeiro de 2016. Tendo como base as necessidades de informações apresentadas no capítulo anterior, foi realizado, no mínimo, um processo de busca para cada pesquisador. A busca simples no acervo da HDB está integrada à da BN Digital. A busca avançada permite as seguintes combinações:

Opção 1

1. Local definido
2. Período definido
3. Periódico definido
4. Texto ou palavras-chave

Opção 2

1. Local definido
2. Período indefinido (Todos)
3. Periódico definido
4. Texto ou palavras-chave

Opção 3

1. Local definido
2. Período definido
3. Periódico indefinido (Todos)
4. Texto ou palavras-chave

Opção 4

1. Local definido
2. Período indefinido (Todos)
3. Periódico indefinido (Todos)
4. Texto ou palavras-chave

Opção 5

1. Período definido
2. Local indefinido (Todos)
3. Periódico indefinido (Todos)
4. Texto ou palavras-chave

Opção 6

1. Período definido
2. Local indefinido (Todos)
3. Periódico definido
4. Texto ou palavras-chave

Opção 7

1. Periódico definido
2. Período indefinido (Todos)
3. Local indefinido (Todos)
4. Texto ou palavras-chave

PESQUISADOR 1

Sobre a busca

Na entrevista o pesquisador informou ter realizado uma busca rápida no acervo da HDB inserindo o nome do sociólogo sobre o qual irá desenvolver a dissertação, escolhendo como local São Paulo e delimitando o período entre as décadas de 1950 e 1980. Obteve três resultados.

Na busca atual optou-se por usar o nome do sociólogo, mas sem delimitar o local, apenas o período, que compreendeu a vida de Hiroshi Saito, de

1919-1983. Adotando esta estratégia, entre as 143 ocorrências, 80 referiam-se ao tema da pesquisa.

O número não excessivo de ocorrências permitiu a leitura de cada resultado.

Comentários sobre os resultados

A partir da década de 1950, a produção acadêmica de Hiroshi Saito começa a aparecer nas páginas de jornal, especialmente em notas sobre publicações acadêmicas. Entre os temas apontados nos jornais estão a aculturação dos japoneses no Brasil e o cooperativismo na região da cidade de Cotia, em São Paulo. Pelas publicações fica-se sabendo também que Saito lecionou na Universidade de Kobe. Em uma reportagem sobre imigrantes japoneses, o sociólogo cede ao jornal fotos realizadas por ele no Japão.

Prêmios recebidos por suas pesquisas, assim como viagens para realizar palestras em outros estados também são registrados nos jornais. Ainda com relação à viagem ao Japão, fica-se sabendo que em Kobe Saito coordenou aulas de português, como escrito no *Diário de Notícias*. A vinda de delegações japonesas recepcionadas pelo professor também viram notícia no *Diário Carioca*.

Pelas notícias dos jornais, toma-se conhecimento que Saito foi um dos encarregados da tradução de *Casa grande e senzala*, de Gilberto Freyre, para o idioma japonês. Ainda na década de 1950, Hiroshi Saito assina um artigo de página inteira sobre a imigração japonesa em *O Semanário* (16-23/05/1957).

Em 1969, o professor é entrevistado para reportagens sobre a integração do nisei no Brasil, tanto no *Diário da Noite* quanto no *Correio da Manhã*. Paralelamente seguem as pequenas notas sobre participação em bancas, palestras, lançamentos de livros, na comissão julgadora de prêmio literário sobre a cultura japonesa. Observa-se ainda que Saito torna-se fonte em diversas notícias, como a do *Jornal do Brasil*, em 1972, que trata sobre os bairros que já foram dominados por imigrantes em São Paulo e, em 1974, sobre a chegada de menor número de imigrantes japoneses ao Brasil.

As notícias mostram a trajetória do sociólogo por instituições ligadas à imigração japonesa, como o Museu da Imigração Japonesa, e os compromissos em diferentes estados brasileiros. Nos anos 1980, em 1º/11/1983, uma nota curta relata a morte do sociólogo aos 64 anos em São Paulo.

O número de ocorrências permitiu uma análise melhor dos conteúdos.

Critérios de busca

Local: Todos (os estados brasileiros)

Periódicos: Todos (os jornais digitalizados que fazem parte do acervo)

Palavra-chave: "Hiroshi Saito"

Períodos:

1920-1929	1960-1969
1930-1939	1970-1979
1940-1949	1980-1989
1950-1959	1990-1999

Resultados

1920-1929: sem resultados.
1930-1939: 21 ocorrências em 10 jornais.
1940-1949: 3 ocorrências em 3 jornais.
1950-1959: 38 ocorrências em 11 jornais.

1960-1969: 16 ocorrências em 8 jornais.
1970-1979: 29 ocorrências em 8 jornais.
1980-1989: 17 ocorrências em 7 jornais.
1990-1999: 19 ocorrências em 5 jornais.

PESQUISADOR 2

Sobre a busca

Ao elaborar o projeto de doutorado, buscou nos jornais (na hemeroteca tradicional da BN do Brasil) a repercussão da crise que levaria à reforma da psiquiatria no país. Como afirmou, a falta de resultados foi uma resposta. Pesquisa agora a assistência psiquiátrica pública do Ministério da Saúde via Serviço Nacional de Doenças Mentais, órgão responsável por políticas públicas psiquiátricas no Brasil nas décadas de 1960 e 1970. Notícias e reportagens de jornais servem de complemento à pesquisa

realizada em periódicos científicos, especialmente tratando da reforma administrativa. O pesquisador buscaria em jornais de São Paulo e do Rio. Apesar da indicação das duas cidades, optou-se por deixar o local de cobertura aberto, pois Brasília já existia e o SNDM poderia ser citado em algum jornal da capital. Como a instituição foi criada em 1941, o período escolhido incluiu de 1940 a 1979.

Comentários sobre os resultados

Em 1941, o *Diário de Notícias* lista a nova organização do Departamento Nacional de Saúde, com a criação de 25 órgãos, entre eles o Serviço Nacional de Doenças Mentais, e sete delegacias. A criação levou à extinção da Divisão de Assistência a Psicopatas do Distrito Federal, passando as suas atribuições ao Serviço Nacional de Doenças Mentais (SNDM).

Os poucos registros sobre o SNDM nos anos 1970 e 1980 são atribuídos ao fato de a instituição ter mudado de nome, passando a se chamar de Divisão Nacional de Saúde Mental na década de 1970, como noticiado pelo *Jornal do Brasil* em 26/05/1970.

Críticas à situação do setor à época aparecem nas páginas do *Correio da Manhã*. Na edição de 05/05/1970 do mesmo *Correio* há uma crítica à situação em que chegou o estado do tratamento de doentes mentais no Brasil. As verbas do SDNM não dariam nem para cobrir as despesas do estado da Guanabara.

Devido à troca de SDNM para Divisão Nacional de Saúde Mental, realizou-se nova busca com este nome nos anos 1970.

Apesar do número expressivo de resultados na década de 1970, os resultados não tratam sobre questões administrativas. Pelas notícias recuperadas, nota-se apenas que os anos 1970 são marcados pela preocupação do aumento do consumo de drogas, o que, segundo os entrevistados, é um problema ligado às doenças mentais.

Como os resultados não foram satisfatórios, decidiu-se empreender mais uma tentativa, com “nova política de saúde mental”, na década de 1970.

No *Jornal do Brasil* (04/10/1972), é mencionada a nova política de saúde mental. Jornais regionais, como o *Diário de Pernambuco* (1980), abordam o que foi realizado por meio da nova política de saúde mental estabelecidas em nível estadual.

No *Jornal do Brasil* (29/01/1974), sai a notícia Saúde Mental já conta com política nacional. Em edição de 1977, porém, o jornal escreve que um núcleo da Secretaria Nacional de Programas Especiais “está elaborando uma política nacional de saúde mental”.

Critérios de busca

Local: Todos	Períodos:
Periódicos: Todos	1940-1959
Palavra-chave: “Serviço Nacional de Doenças Mentais”	1950-1959
	1960-1969
	1970-1979

Resultados

1940-1949: 1071 ocorrências em 30 publicações.	1960-1969: 694 ocorrências em 26 publicações.
1950-1959: 1154 ocorrências em 48 publicações.	1970-1979: 58 ocorrências em 12 publicações.

PESQUISADOR 3

Sobre a busca

O pesquisador já tinha experiência prévia em buscas em jornais, mas não no acervo digital. Havia sido uma consulta muito específica sobre carnaval em um único jornal, a *Folha do Norte*. As buscas serviram de material de apoio à tese de uma professora.

Na presente busca, optou-se por focar no Pará, estado a que pertencia Macapá. Do Amapá, não há jornais na HDB. Especificado o local, os critérios escolhidos foram todos os periódicos em todos os períodos cobertos.

Após não obter resultado significativos com “Fortaleza São José do Macapá”, optou-se por uma busca mais abrangente, com nomes de doenças e o local de interesse.

Comentários sobre os resultados

Realizou-se a busca, mas desde o princípio se tinha noção da dificuldade de completar a consulta com sucesso.

No período coberto pelo pesquisador não havia circulação de jornais no Brasil. Sua pesquisa encerra-se justamente em 1808, ano de criação dos dois primeiros jornais nacionais. Além disso, o local de interesse é o Amapá, que à época pertencia ao Pará. O acervo da HDB não contém publicações do Amapá, nem mesmo de outros períodos mais recentes.

Em *O Jornal*, de 25/09/1900, havia duas notas sobre o tema epidemia de varíola. Uma delas referente a Macapá, informava que o havia regressado da localidade o “desinfectador” da repartição do serviço sanitária, Octavio Jorge, que havia seguido para Macapá a fim de isolar os doentes acometidos de varíola.

O jornal *Republica* (Belém, 14/10/1900) noticia que já está “debellada”, felizmente, a epidemia de varíola que estava grassando em Macapá. Dos 12 doentes acometidos do terrível “morbus, falleceram” 9.

Como os resultados foram pouco satisfatórios, fez-se então uma busca com a palavra-chave “Macapá”. Nos períodos de 1740-1749 e de 1800-1809, não houve ocorrências. De 1810 a 1819, as ocorrências foram baixas, 26 em 4 jornais. Nessas, havia referência à cidade e até mesmo a obras de restauração da fortaleza, mas não a doenças. A combinação da palavra Macapá e sarampo não trouxe resultados.

Critérios de busca

Local: Pará

Períodos:

Periódicos: Todos

Todos

Palavra-chave: “epidemia + varíola + Macapá”

Resultados

121 ocorrências em 12 jornais.

PESQUISADOR 4

Sobre a busca

Pesquisador já realizou diversas buscas bem-sucedidas na HDB nas quais utilizou palavras-chave como “benzedura”, “espinhela caída”. No mestrado, tratou sobre o século XIX, agora dedica-se ao XX. “O material encontrado sempre foi fundamental para contextualizar e encontrar vestígios estimulantes para a compreensão do que estava buscando.”

Nesta busca, tomou-se a decisão de iniciar por um assunto mais amplo, como “plantas medicinais”, por ser

associada às práticas de curas de benzedoras e curandeiros.

Comentários sobre os resultados

Médicos anunciam seus serviços informando que usam remédios feitos somente à base de plantas medicinais brasileiras. Em artigo da *Pacotilha* (Maranhão, 12/03/1912), o médico Monteiro da Silva (o mesmo que costuma anunciar em outros jornais) manda uma indireta para os colegas que não receitam plantas medicinais aos pacientes. “E assim uma flora tão rica em vegetais úteis não tem sido zelada com carinho pela classe médica, que teria bastante motivo de satisfação se receitasse as nossas plantas medicinais e observasse as suas virtudes curativas, sem prejudicar nenhum órgão do organismo” (“Os segredos das Selvas”, p. 1-2).

No jornal *A Noite* (18/05/1918), Medeiros e Albuquerque questiona “medidas sanitárias” de decreto assinado pelo governo. O texto oficial diz: “Ninguém deverá, por nenhuma espécie de anúncio ou de aviso público, tratar ou oferecer para tratar quem quer que seja de moléstias venéreas, nem receitar, nem oferecer-se para receitar nenhum remédio para as referidas moléstias, nem dar nenhuma opinião que se relacione a este tema.” O colunista critica a ordem do governo, mas pondera que isso poderá reduzir também o número de anúncios “charlatanescos”. (p 1).

Nos anos 1920, as plantas medicinais passam a ser citadas como parte das fórmulas dos remédios.

O *Jornal de Recife* (12/04/1931) reproduz partes do que foi publicado pelo governo no *Diário Oficial* de 19/01/1931: “O commercio das plantas medicinaes por atacado é privativo das pharmacias e drogarias, cabendo exclusivamente às primeiras vendas ao publico taes plantas a varejo, quando não sejam tóxicas.” (“O decreto da profissão pharmaceutica e seu exercício no Brasil”, p. 2).

A partir da década de 1930, são raros os anúncios de remédios caseiros, mas há páginas repletas de anúncios de remédios produzidos pela indústria farmacêutica. A edição de 09/07/1931 do *Diário da Noite* traz uma nota sobre o pedido que os proprietários de ervanários querem fazer ao ministro do trabalho para que modifique a lei de que a venda de ervas medicinais só pode ser vendida em farmácia e drogarias. O *Correio de S. Paulo* mostra a prisão de um curandeiro na edição de 08/07/1933.

Critérios de busca

Local: Todos

Períodos:

Periódicos: Todos

1900-1909	1940-1949
1910-1919	1950-1959
1920-1929	1960-1969
1930-1939	

Palavra-chave: “plantas medicinais”

Resultados

1900-1909: 8 ocorrências em 8 jornais.
1910-1919: 48 ocorrências em 13 jornais.
1920-1929: 51 ocorrências em 22 jornais.
1930-1939: 377 ocorrências em 43 jornais.
1940-1949: 2819 ocorrências

em mais de 50 jornais.
1950-1959: 2362 ocorrências em mais de 50 jornais.
1960-1969: 456 ocorrências em mais de 35 jornais.

PESQUISADOR 4

Sobre a busca

Em seguida, optou-se por escolher uma doença relacionada a benzeduras. Por ter sido citada pelo pesquisador, escolheu-se “espinhela caída”, como era chamado o reumatismo no início do século XX.

Comentários sobre os resultados

A maioria dos resultados remete a anúncios de benzedoras, que curam não só espinhela caída, mas cobreiros, mau olhado, “quebrantos” e outros males. De modo geral, são anúncios de senhoras. Quando são citadas em notícias, são chamadas de benzedoras, rezadeiras e até mesmo de feitiçeras. O tom das notas jornalísticas nunca é positivo, como a que saiu no jornal *A Notícia* (21/12/1913): “Bruxedo e mau olhado”.

Na década de 1910, há menos ocorrências, mas o tom de reprovação persiste nos artigos jornalísticos. Nos anos 1920, curas e benzeduras são ensinadas n’*O livro da bruxa ou Manual da Cartomante*, “o trabalho mais completo que se tem publicado até hoje”, como anunciado no *Jornal do Brasil* (17/01/1924). Notícia também é a prisão de curandeiros, como a da Mãe D’Água em 27/02/1920 no *Correio da Manhã*.

Nos jornais das décadas de 1930 e 1940 aparecem folhetins com personagens que contraem ou curam doenças como a espinhela caída e outros males. As feitiçarias são tema de textos assinados por médicos, que abordam os perigos de recorrer a práticas não científicas.

Critérios de busca

Local: Todos	Períodos:	
Periódicos: Todos	1900-1909	1940-1949
	1910-1919	1950-1959
	1920-1929	1960-1969
Palavra-chave: “espinhela caída”	1930-1939	

Resultados

1900-1909: 111 ocorrências em 10 jornais.
1910-1919: 29 ocorrências em 18 jornais.
1920-1929: 80 ocorrências em 28 jornais.

1930-1939: 56 ocorrências em 25 jornais.
1940-1949: 34 ocorrências em mais de 23 jornais.

PESQUISADOR 5

Sobre a busca

O pesquisador tem experiência em busca no acervo da hemeroteca da Biblioteca de Lisboa, onde obteve um artigo de seu interesse na *Gazeta de Lisboa*.

Do acervo brasileiro ele pesquisou somente o jornal *O Patriota*, não na versão em papel tampouco no acervo *on-line*, mas em um cópia em CD. Em suas pesquisas iniciais buscou por nomes de instituições e de personagens da história.

Agora pretende demonstrar como cirurgiões atuaram diante de vínculos

com as forças armadas e as câmaras municipais coloniais.

Comentários sobre os resultados

Em *O Universal* (11/06/1830) há uma indicação de que os médicos devem retirar suas licenças na Câmara Municipal. Uma das polêmicas da época é a prática da medicina por estrangeiros.

A busca recupera as duas palavras, mas não quer dizer que elas estejam relacionadas.

As buscas não foram frutíferas, mostrando que sem entender o contexto ou o conteúdo tornar-se difícil seguir adiante e recuperar informações que realmente serão úteis.

Critérios de busca

Local: Minas Gerais

Periódicos: Todos

Palavra-chave: “cirurgiões + câmara”

Períodos:

1820-1829

Resultados

1820-1829:
9 ocorrências em 5 jornais

PESQUISADOR 6

Sobre a busca

O período de cobertura do estudo é anterior à existência de jornais no Brasil. Mesmo assim, decidiu-se por pesquisar algumas décadas após a produção de publicações no país.

Após tentativas frustradas, optou-se por termos mais amplos, como sangradores, associado a quem realizada o serviço à época, escravos treinados, comumente chamados de “pretos”.

Comentários sobre os resultados

Na edição do *Correio Braziliense* é citado que “(...) a bordo desses navios não poderão andar cirurgiões para curar os escravos, pelos não haver, ou por outra razão equivalente, serão os donos ou fretadores obrigados a trazer a bordo dos ditos navios pretos sangradores, inteligente e experimentados no tratamento das moléstias, de que ordinariamente são infectados os ditos escravos, e no conhecimento dos remédios próprios e adequados, de que eles usam em seus curativos.

Nas décadas seguintes, a palavra sangrador quando associada a um negro referia-se a leilões de escravos, inclusive daqueles que tinham conhecimento do ofício de barbeiro e sangrador. Estas informações são encontradas em anúncios, mas não em reportagens.

Critérios de busca

Local: Todos

Periódicos: Todos

Palavra-chave: “sangradores + pretos”

Períodos:

1810-1819	1840-1849
1820-1829	1850-1859
1830-1839	1860-1869

Resultados

1810-1819: 1 ocorrência em 1 jornal
1820-1829: 1 ocorrência em 1 jornal
1830-1839: 10 ocorrências em 3 jornais

1840-1849: 4 ocorrências em 2 jornais
1850-1859: 2 ocorrências em 2 jornais
1860-1869: 1 ocorrências em 1 jornais

PESQUISADOR 7

Sobre a busca

Pesquisador já realizou buscas prévias tanto no acervo físico quanto digital da HDB. Ao realizar a monografia, buscou revistas ilustradas no período 1928-1933.

Atualmente, para a dissertação, busca anúncios ou notícias em jornais, no período 1820-1870, a respeito das parteiras. Em suas consultas recuperou anúncios, notícias e crônicas.

Baseada na experiência bem-sucedida da pesquisadora adotou-se a palavra-chave “parteira” e o período de sua pesquisa. Após as primeiras

buscas observou-se que ama de leite sempre era mencionada nos anúncios das profissionais instaladas no Rio de Janeiro.

Os estados, Bahia e Rio de Janeiro, escolhidos deram-se pelo fato de serem estes dois os primeiros a contar com escolas de medicina e da exigência de um certificado para atuação como parteira.

Comentários sobre os resultados

O maior número de ocorrência refere-se a parteiras que publicam anúncios para divulgar seus serviços. Quando anunciam, nota-se o cuidado em dizer onde a parteira obteve sua formação e/ou certificação. Comum é a menção à maternidade de Paris, especialmente por aqueles que assinam com nomes franceses.

O jornal também era o meio de os pais agradecerem por um trabalho bem feito, como o registrado no *Diário do Rio de Janeiro* (20/03/1932).

A *Revista de Medicina Medica* apresenta diversas histórias envolvendo parteiras, em muitas delas médicos acusando as mulheres de imperícia. Na coleção da Biblioteca Nacional há ainda outras publicações redigidas por médicos (*Annaes de Medicina Brasiliense, Revista Medica Brasiliense*).

No *Diário do Rio de Janeiro* (10/02/1942) há um alerta a quem não apresentar a carta que autoriza a prática de boticários, sangradores, parteiras etc.

Em outras edições nas décadas seguintes, parece haver um grande interesse em coibir a atuação daquelas sem licenças, seja por meio de notas jornalísticas ou anúncios. Nos anos 1860 há anúncios com recomendações de parteiras mais antigas às mais novas.

Critérios de busca

Local: Rio de Janeiro / Bahia

Períodos:

Periódicos: Todos

1820-1829 1850-1859

1830-1839 1860-1869

1840-1849

Palavra-chave: “parteira”

Resultados

1820-1829: 23 ocorrências em 5 jornais
1830-1839: 212 ocorrências em 21 jornais
1840-1849: 448 ocorrências em 16 jornais

1850-1859: 651 ocorrências em 19 jornais
1860-1869: 1309 ocorrências em 27 jornais

PESQUISADOR 8

Sobre a busca

Os manuais de medicina que serão tema da dissertação do pesquisador foram escritos antes da impressão de jornais no Brasil.

Por meio das buscas, procurou-se saber se mesmo décadas depois do lançamento eles ainda eram citados em notícias.

Na primeira busca empreendida optou-se por usar o nome do manual mais conhecido, o *Erário Mineral*,

lançado em 1735, de autoria de Luiz Gomes Ferreira.

O nome do autor, por apresentar muito homônimos não se mostrou como uma boa escolha.

Comentários sobre os resultados

O *Erário Mineral* aparece em diversos anúncios de livrarias, seja de interessados em vendê-lo quanto de interessados em comprá-lo. Na *Revista Médica Fluminense*, publicada pela Academia Imperial de Medicina do Rio de Janeiro em 1838, em que é mencionado em uma reportagem.

No *Archivo Medico Brasileiro* (1846), em uma crônica que manifesta o interesse crescente da população por obras como o *Erário Mineral* de Ferreira. Em outra de *O Mercantil*, o autor de uma crônica trata o *Erário Mineral* como um livro já esquecido. Ainda em 1949, a leitura do *Erário* é recomendada em reportagem do *Jornal do Brasil*.

Artigo sobre a Medicina Brasileira, no *Jornal do Brasil* (1957), cita o *Erário* como sendo “de fato e de direito, o primeiro compêndio de medicina brasileira escrito em vernáculo”. Reportagem do *Correio da Manhã* (Minas Gerais, 1955) fala sobre a importância do livro.

Critérios de busca

Local: Todos

Periódicos: Todos

Palavra-chave: “erário mineral”

Períodos:

1820-1829	1890-1899
1830-1839	1900-1909
1840-1849	1910-1919
1850-1859	1920-1929
1860-1869	1930-1939
1870-1879	1940-1949
1880-1889	1950-1959

Resultados

1820-1829 / 1830-1839: Sem resultados
 1840-1849: 4 ocorrências em 3 publicações
 1850-1859: 7 ocorrências em 6 jornais
 1860-1869: 4 ocorrências em 3 jornais
 1870-1879: 1 ocorrências em 1 jornal
 1880-1889: 1 ocorrências em 1 jornal
 1890-1899: 5 ocorrências em 3 jornais

1900-1909: 1 ocorrência em 1 jornal
 1910-1919: 4 ocorrências em 4 jornais
 1920-1929: sem resultados
 1930-1939: 3 ocorrências em 2 jornais
 1940-1949: 2 ocorrências em 2 jornais
 1950-1959: 3 ocorrências em 3 jornais

PESQUISADOR 9

Sobre a busca

A atuação do Instituto Behring de Terapêutica Experimental (IBTE) é o foco do trabalho de pós-doutorado da pesquisadora, especificamente os desdobramentos da Segunda Guerra Mundial, quando diretores foram presos a mando de Getúlio Vargas.

O pesquisador também busca notícias de usos dos produtos. Por isso, optou-se por pesquisar as décadas de 1930-1939 e 1940-1949.

Critérios de busca

Local: Todos

Periódicos: Todos

Palavra-chave: "Instituto Behring"

Períodos:

1930-1939

1940-1949

Resultados

1930-1939: 15 ocorrências em 7 jornais
1940-1949: 15 ocorrências em 8 jornais

Comentários sobre os resultados

Os jornais da década de 1930 fazem menção aos impostos pagos pela companhia e quais vacinas são produzidas pelo Behring.

A história de Emil Von Behring, de origem alemã, é contada na edição de 6 de fevereiro de 1938 de *O Imparcial*. Novamente em 19 de dezembro de 1940, o jornal publica uma matéria sobre von Behring, assim como a *Gazeta de Notícias* (04/12/1940) também fala da vacina criada pelo cientista. No jornal *O Imparcial* (24/06/1947), o IBTE aparece entre as firmas brasileiras que estavam na lista negra dos Estados Unidos da América.

Na década de 1940, a *Gazeta de Notícias* (19/10/45), noticia que todas as patentes de invenção, modelos de utilidade, desenhos ou modelos industriais, marcas de indústria de comércio concedidas pelo Departamento Nacional da Propriedade Industrial que tenham sido usados pelas firmas Merck, Bayer e Behring, ora sob regime de liquidação, ficam doados à Fundação Brasil Central.

PESQUISADOR 10

Sobre a busca

O grupo modernista chamado verde-amarelo é o objeto de estudo do pesquisador. Este grupo era formado por Menotti del Picchia (1892-1988), Plínio Salgado (1895-1988), Guilherme de Almeida (1890-1969) e Cassiano Ricardo (1895-1974).

Foram realizadas pelo menos quatro buscas, sempre no período da década de 1920. Uma com o nome do grupo e as demais pelos nomes dos protago-

nistas, optando-se por ampliar para todos os estados brasileiros, não apenas SP, apesar de os resultados serem em maior número nos jornais paulistas.

Critérios de busca

Local: Todos

Periódicos: Todos

Palavra-chave: “grupo verde-amarelo”

Período:

1920-1929

Resultados

1920-1929:
26 ocorrências em 11 jornais
Busca por:
“verde + amarelo + Cassiano Ricardo”: 64
ocorrências em 18 jornais
Apenas “Cassiano Ricardo”: 841 ocorrências
em mais de 50 jornais.

“verde + amarelo + menotti del picchia” 28
ocorrências em 12 jornais.
Apenas “menotti del picchia” 1794 ocorrên-
cias em mais de 50 jornais.
“verde + amarelo + Plínio Salgado” 28 ocor-
rências em 10 jornais.
Apenas “Plínio Salgado” 1257 ocorrências
em 46 jornais.

Comentários sobre os resultados

Nota do jornal *O Paiz* cita um manifesto assinado por Cassiano Ricardo e Menotti Del Picchia no jornal *Correio Paulistano* de 17/05/1929 sobre o movimento verde e amarelo.

É nas colunas que os envolvidos na rixa “debatem”.

No *Correio Paulistano* de 26/09/1927 lê-se uma carta de Oswald de Andrade encaminhada a Menotti del Picchia sobre determinada poesia. Na edição de 11/01/1927, p. 4, enorme artigo falta sobre o movimento. Cassiano Ricardo (06/07/1926) e Plínio assinam artigos no mesmo jornal sobre o movimento (04/06/1927 / 10/06/1928).

Um texto de Abner Mourão analisa a disputa entre o Pau Brasil (Oswald de Andrade e cia) e o Verde e Amarelo (Cassiano Ricardo e cia.) na edição de *O Paiz* (30/09/1925).

Artigo da publicação *A.B.C.* (21/05/1927) explica o que é o verde-amarellismo.

Menotti del Picchia assinou diversas colunas chamadas *Verde e Amarelo* no *Correio Paulistano* (01/11/1924 - 23/09/1925). Na coluna *Nossa orientação* (20/01/1926), Menotti explica a orientação tomada por ele, Cassiano Ricardo e Plínio Salgado. Por várias vezes, Menotti é citado em colunas dos parceiros de movimento. Nessa década, Menotti é eleito deputado estadual de São Paulo.

Juca Branco assina uma sarcástica coluna sobre Menotti no *Diário Nacional* (06/08/1927).

Com o nome dos três principais atores juntos, “cassiano ricardo + menotti del picchia + plínio salgado”: 52 ocorrências em 10 jornais, mas sempre com mais ocorrências no *Correio Paulistano*.

PESQUISADOR 11

Sobre a busca

A busca mostrou-se muito difícil uma vez que o próprio pesquisador ainda não tem clareza do que busca. Ele pretende encontrar documentos referentes à atuação do Hospital São Francisco de Belo Horizonte, em relação ao atendimento público de saúde prestado aos desempregados durante o governo Vargas até 1988, com a implantação do Sistema Único de Saúde.

Comentários sobre os resultados

Grande parte das ocorrências fazia referência ao Hospital São Francisco de Assis no Rio de Janeiro, especialmente anúncios de médicos que foram residentes no hospital da então capital.

A busca não foi bem-sucedida. Faltaram informações para um refinamento melhor da necessidade do pesquisador.

Critérios de busca

Local: Minas Gerais

Periódicos: Todos

Palavra-chave: "Hospital São Francisco"

Períodos:

Todos

Resultados

71 ocorrências em 6 jornais.

PESQUISADOR 12

Sobre a busca

A história da psicanálise no Rio de Janeiro nas primeiras décadas do século XX foi tema da tese do pesquisador, que continua a busca informações sobre o assunto.

O pesquisador já teve a experiência na Hemeroteca Digital brasileira, onde buscou por nomes dos autores, como Juliano Moreira, Henrique Roxo, Julio Porto-Carrero.

A busca baseou-se na experiência

realizada pelo pesquisador, utilizando a palavra “psychanalyse”.

Comentários sobre os resultados

Em matéria de *O Paiz* (05/12/1925) são citados os nomes de médicos conhecidos no Rio de Janeiro por tratar alienados, como Juliano Moreira, Henrique Roxo, Franco da Rocha, Medeiros e Albuquerque, Renato Jardim.

Na edição de 23/03/1928 de *O Paiz* é anunciado um curso de psicanálise realizado pela Associação Brasileira de Educação. Os professores eram Julio Porto Carrero e Deodato de Moraes. O curso foi notícia durante meses (pequenas notas sobre as palestras, onde está sendo realizado, a ótima assistência).

Na edição de 07/07/1928, membro da Sociedade Brasileira de Psicanálise de São Paulo participam de reunião junto aos cariocas para a criação de uma unidade da Sociedade também no Rio de Janeiro. A reunião noticiada também em *O Jornal* (06/07/1928).

A “psychanalyse perante a Sciencia”, de Pedro Estevão José Oliva (*O Jornal*, 17/06/1928) é uma crítica ao modo como a disciplina esta se disseminando na cidade, não por meio de profissionais, mas de forma vulgar, por qualquer um. (menção ao fato do curso de psychanalyse ser aberto à população.

Inicialmente, a psicanálise aparecia como interpretação de sonhos.

Com a grafia “psicanálise”, local RJ e década de 1920-1929 começam a aparecer resultados apenas a partir da década de 1920.

Nos anos 1930, senhoras que trabalham com quiromancia usam a palavra em seus anúncios. A psicanálise aparece em alguns anúncios ao lado de tarô, astrologia e quiromancia, como nos anúncios do Dr. Florial veiculados em diferentes edições do *Jornal do Brasil* nos anos 1940.

O Malho (15/03/1936, p.32) faz piada tendo como inspiração a psicanálise (Curiosidades da psicanálise)

Muitos anúncios de palestras e tratamentos, especialmente assinados pelo dr. Gastão Pereira da Silva, que escreve inclusive um artigo explicando a área de atuação da psicanálise na edição de 19/01/1941 de *A Noite*.

Critérios de busca

Local: Rio de Janeiro

Periódicos: Todos

Palavra-chave: “psychanalyse”

Períodos:

1900-1909

1910-1919

1920-1929

1930-1939

1940-1949

Resultados

1900-1909:
Sem resultados
1910-1919:
15 ocorrências em 8 jornais
1920-1929:
254 ocorrências em 22 jornais
1930-1939:

1270 ocorrências em 38 jornais
1940-1949:
70 ocorrências em 19 jornais
1950-1959:
9 ocorrências em 4 jornais

PESQUISADOR 13

Sobre a busca

Notícias sobre a assistência prestada a alienados no estado do Mato Grosso, entre 1931 e 1957, são o foco desta pesquisa. O pesquisador já realizou buscas em jornais e revistas antigos pertencentes a coleções físicas. O interesse à época foram as epidemias de febre amarela.

Critérios de busca

Local: Mato Grosso

Periódicos: Todos

Palavra-chave: “alienados”

Períodos:

Todos

Resultados

100 ocorrências em 31 jornais

Comentários sobre os resultados

Em *A Cruz* (01/08/1915) há a notícia da inauguração do Pavilhão de Alienados na Santa Casa de Misericórdia, também veiculada em *O Matto-Grosso* (08/08/1915). Curiosamente, outro pavilhão seria construído anos mais tarde. *O Matto-Grosso* noticia o lançamento da pedra fundamental (08/04/1920). No mesmo jornal (27/01/1929), uma notícia chama a atenção em 1929. O pavilhão construído para abrigar os alienados teve seu destino alterado pelo governador da época, que ali decidiu inaugurar uma enfermaria para mulheres, crianças e a maternidade da Santa Casa de Misericórdia. Decisão não muito bem vista pelos autores do texto, que ao final do governo comemoram que o governador está fora do poder.

As notícias recuperadas mostram o quanto a palavra “alienados” aparece ligada ao espiritismo naquela comunidade, sendo o espiritismo visto com bastante preconceito. “O caminho mais próximo do hospício é espiritismo.” (*A Cruz*, 15/07/1917, p. 2). Notou-se o uso frequente de asylo de alienados e menção ao asylo de alienados do Coxipó. Porém, as buscas por essas palavras-chave não renderam resultados significativos.

Nas primeiras décadas de 1900, a Santa Casa de Misericórdia é a única opção de tratamento a alienados no MT. Na edição de 1909 do jornal *O Pharol* a uma nota sobre haver 34 doentes na Santa Casa, sendo 13 homens, 21 mulheres, dos quais 11 alienados.

Na edição de 11/05/1899, a *Gazeta Oficial* coloca em um artigo a necessidade de uma lei que regule a situação dos alienados no estado. Os textos recuperados são pouco jornalísticos, mostrando mais opinião.

A busca por “doença mental” resultou apenas 5 ocorrências em 2 jornais. Nenhum dos resultados estava relacionado ao tema da pesquisa. Ao utilizar a palavra “louco”, notou-se que os resultados traziam a palavra ligada de modo mais figurativo do que a doentes mentais, muitas vezes presente em folhetins, mais relacionada a pessoas que tomam atitudes não consideradas normais, mas sem terem necessariamente problemas de saúde mental.

9 INTERPRETAÇÃO DOS RESULTADOS

Como recuperar uma informação se ela não estiver representada? O que significa representar uma informação? Que informação é esta representada? Como seria a representação ideal de uma coleção de jornais e revistas? Como poderia ser possível melhorar o acesso do usuário à coleção de publicações periódicas? Guiada por tais indagações, iniciou-se a revisão bibliográfica.

Primeiro, veio a compreensão do tipo de informação de que trata a Ciência da Informação, então as correntes de pensamentos que ora agregam, ora separam organização do conhecimento e organização da informação. No esquema escolhido, as duas atividades apresentam-se separadas, cada qual com seus objetivos e atribuições.

Enquanto a organização do conhecimento reflete sobre fundamentos teóricos, metodologias e sistemas, a organização da informação volta-se para a prática, a aplicação do que a primeira realiza. A OI não poderia existir sem a OC. Tanto em uma quanto na outra está o desejo de representar “o mundo” para que este seja recuperado pelo usuário. Para que representar a informação de forma temática ou descritiva se não para que seja localizada de forma mais precisa por aqueles que dela necessitam?

Há mais de 100 anos, profissionais, sejam bibliotecários, sejam cientistas da informação, debruçam-se sobre o desafio de encontrar soluções para o armazenamento de um volume crescente de informações. Classificação, catalogação, indexação foram algumas das contribuições mais importantes, mas não as únicas ou as últimas. O desafio da organização de itens, da recuperação da informação que melhor atenda às necessidades dos usuários renova-se constantemente.

As tecnologias utilizadas nesta segunda década do Século XXI incentivam o sonhar grande. Afinal, quem poderia imaginar um repositório que reunisse boa parte da informação de que precisamos para facilitar nossas vidas – como havia pensado Otlet há um século?

Se hoje ainda não existem mecanismos que respondam a todas as perguntas dos usuários com precisão, sabe-se que houve um considerável avanço. As hemerotecas são um exemplo disso. Primeiro bastava inserir algumas poucas informações sobre os jornais em um catálogo, classificá-los para que fossem encontrados nas estantes e guardá-los da melhor maneira possível. Com o avanço dos sistemas eletrônicos, o modelo tão bem estabelecido começou a se tornar insuficiente. Havia um novo fato, os jornais se deterioravam pelo manuseio.

E não apenas isso. Mesmo que ficassem para sempre intactos nas estantes, estariam ameaçados. Sabe-se que o papel jornal é um tipo inferior de papel, de baixo custo, feito para ser tão efêmero quanto um jornal diário. Tão logo esta fragilidade foi percebida, as Bibliotecas Nacionais iniciaram seus planos de microfilmagem, tentativa de garantir que aquele retrato da sociedade não desaparecesse.

Como escreve Gleick (2013, p. 20), uma tecnologia, um modelo ou um plano não se torna absolutamente obsoleto, mas serve de base para o que está por vir. Ao percorrer a evolução das ferramentas empregadas em unidades de informação, observa-se que mecanismos de tratamento da informação foram sendo substituídos apenas quando os novos instrumentos atendiam a necessidade daquele momento de forma mais adequada. A própria indexação manual ou mesmo o microfilme chegaram a ser considerados obsoletos com o surgimento de novas tecnologias, para algum tempo depois serem resgatados, tendo seu valor reconhecido.

A microfilmagem – e depois a digitalização – permitiu ao público acesso a uma parte considerável do acervo sem que se precisasse ir pessoalmente até este conjunto de jornais e revistas. O processo de seleção e preparação dos originais para a digitalização possibilitou a “reabilitação” de diversas coleções. Tanto por serem transferidas para uma cópia digital, mas ainda mais por tirá-las do esquecimento/deterioração nas estantes das BN. Parte significativa das publicações, que havia permanecido intocada por anos ou até mesmo décadas, foi limpa, desamassada e recuperada. A Biblioteca Britânica, por exemplo, combinou o processo de digitalização com a construção de um novo armazém, planejado com uma infraestrutura altamente tecnológica para guardar coleção tão frágil como esta impressa no papel jornal.

Durante a revisão de literatura, procurou-se também resgatar o outro tipo de informação ligado a este trabalho, a jornalística. O que é o jornalismo senão uma coleção de notícias, o retrato de determinado momento. É nesta característica que reside o encanto da área. Mesmo com todas as considerações sobre a dificuldade de se manter a imparcialidade ou sobre ideologias e tendências por trás da escolha das notícias, nada diminui o valor do registro jornalístico.

A linguagem jornalística, baseada na objetividade, concisão, ordem direta e simplicidade, contribui para a democratização da informação, para aumentar o alcance das notícias. O texto jornalístico atende do estudante primário ao mais letrado dos leitores. A universalidade dos jornais apresenta-se na diversidade de temas e

abordagens, assim como na variedade do público-alvo. A partir das informações veiculadas nas páginas dos jornais, espera-se que o leitor tenha discernimento, competência em informação, para tirar suas próprias conclusões. Autonomia semelhante na interpretação dos fatos reportados espera-se dos pesquisadores que decidem tomar jornais e revistas como fonte de informação.

Ao ser iniciada a procura de hemerotecas digitais mantidas por Bibliotecas Nacionais mundo afora, havia a expectativa de se encontrar um número considerável. À medida que as buscas foram sendo empreendidas, percebeu-se que apesar de boa parte das Bibliotecas Nacionais já terem começado a digitalizar suas coleções, ainda é pequeno o número de bibliotecas digitais ou de hemerotecas digitais estruturadas – seja por falta de recursos financeiros ou tecnológicos.

Entre as hemerotecas digitais selecionadas, por exemplo, há dois casos em que se percebe que o trabalho já foi iniciado, mas estas ainda não chegaram ao patamar de desenvolvimento das demais. As hemerotecas digitais da Argentina e da Costa Rica já têm um número notável de jornais e revistas digitalizados e catalogados, mas não implementaram o uso de OCR, o que prejudica as buscas no conteúdo. Como descrito anteriormente, a hemeroteca da Costa Rica já adotou filtros que colaboram no refinamento das buscas, mas no caso da coleção argentina só é possível navegar pelos títulos das publicações, tendo o consulente que folhear as edições de interesse para encontrar o que se busca – seguindo, de certa forma, o modelo tradicional de busca da informação nas coleções em papel.

Sobre a escolha das hemerotecas tratadas neste trabalho, sabe-se que não foram incluídas hemerotecas digitais importantes, como as desenvolvidas pelos países europeus nórdicos. Porém, diante da perspectiva de não poder realizar buscas nos acervos por limitações no domínio dos idiomas de origem, optou-se por focar aquelas hemerotecas que não ofereciam este tipo de dificuldade. Acredita-se que com as hemerotecas digitais selecionadas foi possível apresentar um recorte significativo do cenário desse tipo de coleção.

As diferenças de posicionamento das BN podem ser notadas já no acesso às coleções na “Web”. Áustria, Espanha, Estados Unidos, México, Nova Zelândia e Reino Unido permitem a entrada no site por meio de um endereço eletrônico próprio, sem necessidade de navegar pela página inicial da respectiva Biblioteca Nacional ou biblioteca digital.

Tal decisão leva a considerar a hemeroteca como uma unidade de informação autônoma, como um serviço dotado de independência, não apenas como a parte de um projeto (de preservação do patrimônio nacional), o que continuará sendo por estar ligada à BN. Além disso, de forma prática, oferece aos usuários agilidade no acesso. O *link* direto facilita o acesso àqueles usuários pouco familiarizados com o site da(s) BN(s), fazendo com que parte desse público não desista de chegar à parte do site em que está instalada a hemeroteca.

Bélgica e Costa Rica apresentam suas hemerotecas como uma das coleções da Biblioteca Nacional, mas ambas possuem um mecanismo de busca próprio. Como referido anteriormente, a necessidade de navegar pelo site principal da BN aumenta o risco de esse usuário desistir da busca caso encontre dificuldades em encontrar o caminho certo.

Argentina, Austrália, França e Liechtenstein oferecem acesso à coleção digitalizada de jornais e revistas via página inicial de suas respectivas bibliotecas digitais: Trapalanda, Trove, Gallica e eLiechtensteinesia.

Com exceção da Gallica, todas têm um mecanismo de busca específico para a seção de jornais e revistas. No caso francês, mesmo na página específica de publicações periódicas, o usuário realiza as buscas no mecanismo geral da biblioteca digital, seja de maneira simples ou avançada.

A vantagem é ter acesso a todo o acervo, obtendo resultados nos mais diferentes formatos, não se limitando somente à coleção de publicações periódicas. Por outro lado, o processo de busca da informação requer maior atenção e pode gerar frustração quando da obtenção de documentos que não atendem ao que se esperava.

A Hemeroteca Digital Brasileira também é acessada desta forma, via BNDigital. Tal posicionamento valoriza a biblioteca digital e reforça o seu papel de agregadora do conteúdo digital de uma nação. Os jornais e revistas digitalizados não são uma coleção à parte, integram um patrimônio e estão guardados no acervo da instituição responsável pela memória bibliográfica do país.

Em relação às buscas na HDB, as consultas simples são realizadas no mecanismo geral da BNDigital. Para a busca avançada no acervo jornalístico, há um mecanismo específico, no qual o usuário realizará a busca selecionando local, período ou título e associando a este parâmetro uma ou mais palavras-chave. A variedade de mecanismos de busca podem causar confusão aos usuários iniciantes, mas uma vez

familiarizados com as diferentes opções o uso torna-se intuitivo.

Recurso inovador e responsável por um dos maiores avanços no cenário das coleções hemerográficas é o *software* para reconhecimento óptico de caracteres, que permite buscas no conteúdo. A partir do momento em que se passou a usar OCR, o usuário não precisou mais percorrer as páginas dos jornais manualmente, uma a uma, à procura de textos que atendessem a sua necessidade. O recurso, com grau considerável de precisão, seleciona para este usuário as páginas em que se encontram as palavras-chave ou frase desejada.

O uso do OCR não elimina o refinamento que só um humano é capaz de realizar tampouco reduz a necessidade de se refletir sobre os melhores parâmetros e palavras-chave a serem usados, mas faz com que o usuário economize tempo e tenha um alcance muito maior. Sem esse recurso, quantas horas levaria um pesquisador para percorrer as 10 milhões de páginas já digitalizadas pela Biblioteca Nacional brasileira, por exemplo? Atualmente em poucos minutos o mecanismo de busca retorna uma lista de jornais com as páginas que registram a palavra-chave desejada.

Diferente da maior parte das hemerotecas estudadas, os resultados da HDB são apresentados em uma lista composta pelos títulos do jornais e uma indicação numérica à direita de quantas ocorrências foram detectadas das palavras-chave empregadas. Nos demais itens consultados, mas sem resultados compatíveis à busca, lê-se o número zero (Figura 68). Tal apresentação ajuda o usuário a ter uma visão completa dos resultados, já tendo noção de quais veículos abordaram, a princípio, o tema da busca.

A partir desta lista, o usuário deverá clicar no nome do veículo de sua escolha. Imediatamente será conduzido à primeira página do material recuperado. O sistema permite o deslocamento pelas páginas.

Figura 68 – Apresentação dos resultados na Hemeroteca Digital Brasileira

Descrição	Ocorrências ▼
A Republica - 1888 a 1930 - PR_SPR_00214_215554	60
Diário do Paraná: Orgao dos Diários Associados - 1955 a 1983 - PR_SPR_02341 -	8
Correio do Paraná - 1932 a 1965 - PR_SPR_02477_001_171395	3
O Estado do Paraná - 1925 - PR_SPR_02486	2
Ultima Hora - PR_SPR_02487_001_830348	2
O Estado do Paraná - 1925 a 1926 - PR_SPR_02491	1
PR_SPR_02340 - Diário do Paraná	0
A Notícia - 1905 a 1908 - PR_SPR_00522_187666	0
Almanach do Paraná - 1896 a 1929 - PR_SPR_00468_214752	0
Electra - 1901 a 1903 - PR_SOR_00649_001_223778	0

Fonte: Biblioteca Nacional Digital. Disponível em: <<http://bndigital.bn.br/hemeroteca-digital>>. Acesso em: 10 fev. 2016.

Com exceção das já mencionadas coleções da Argentina e da Costa Rica, na comparação entre as demais hemerotecas digitais inseridas neste estudo percebe-se um equilíbrio no desenvolvimento de ferramentas e mesmo na maneira de organizar as coleções. Observa-se o uso de OCR como forma de substituir uma inviável indexação de conteúdos (pelo volume, custo e tempo), assim como a adoção de uma catalogação com o máximo de metadados possível, nem sempre usados no momento das buscas, mas para uma posterior filtragem dos resultados.

Por se tratarem de sistemas relativamente recentes, acredita-se que melhorias possam ocorrer em um futuro próximo, até mesmo inspiradas pela experiência de outras hemerotecas. Oferecer ao usuário mais de uma opção de filtragem, em mais de um momento da busca, torna o trabalho muito mais ágil e com maiores chances de obter resultados que melhor atenderão à necessidade de informação. Ou seja, tendo a opção de filtrar os resultados em mais de uma etapa, o usuário não precisaria repetir todo o processo desde o começo ao ter novos *insights* provenientes da experiência recém-realizada.

A HDB mostra-se em nível de desenvolvimento semelhante ao das demais no que diz respeito à busca e recuperação de conteúdos, apesar da impossibilidade de fazer nova busca a partir do material já recuperado. O mecanismo de busca apresenta os três parâmetros encontrados em praticamente todas as hemerotecas: local, período e título, oferecendo ainda a possibilidade de combinar essas informações com palavras-chave.

Uma ressalva que pode ser feita em relação à busca é a impossibilidade de pesquisar em todos os títulos de todos os estados em todos os períodos disponíveis. É

sempre necessário selecionar pelo menos um período ou um local ou um título. Por outro lado, a busca em todo o acervo, poderia trazer um número tão elevado de resultados que impossibilitaria a revisão do que foi encontrado, especialmente por não ser possível filtrar os resultados em uma segunda etapa da busca, como ocorre em outras hemerotecas pesquisadas.

A adoção de OCR mostra que a HDB está alinhada ao que há de mais tecnológico, por ora, na área. A HDB informa utilizar o “Inteligenciamento DocPro”, processo que envolve a pesquisa por aproximação visual, característica principal da tecnologia DocPro, na qual não são guardadas as palavras exatas e sim a aproximação visual de cada uma. Como explica Bettencourt (2013, p. 7), desta forma as falhas que normalmente acontecem em um OCR comum são minimizadas, o que se traduz em uma taxa de acerto muito superior.

No aspecto da apresentação dos resultados, a HDB assemelha-se à *BelgicaPress*, da Biblioteca Real da Bélgica, que fornece a lista dos jornais e o número de páginas recuperadas. Não é a forma mais completa de apresentar os resultados e esta poderia ser uma das melhorias a serem consideradas. Hemerotecas como a do Reino Unido, do México, da Austrália ou mesmo a da Áustria desenvolveram modelos que oferecem os resultados de forma mais completa ao usuário, além de disponibilizarem diferentes filtros.

Ainda em relação aos resultados, outros aperfeiçoamentos possíveis à HDB seriam a opção de copiar o texto produzido por OCR e uma indicação de quantas ocorrências foram visualizadas. Este parâmetro seria útil no caso de um número elevado de resultados. Outra sugestão seria o armazenamento da busca, para que o pesquisador pudesse retornar no ponto onde parou, sem precisar refazer todo o processo.

Tal procedimento é oferecido, por exemplo, pelas hemerotecas da Austrália e do Reino Unido, nas quais o usuário pode salvar as buscas empreendidas. Outro exemplo de boa prática, que extrapola o universo das hemerotecas, pode ser observado no site Pubmed, em que um sistema de alerta personalizado e customizado envia por e-mail novos resultados das buscas armazenadas. Numa hemeroteca, poderia ser útil em função da progressiva digitalização de novos documentos.

A partir das buscas empreendidas na HDB, notou-se se tratar de um mecanismo de busca bastante intuitivo. Após receber os resultados, o usuário possivelmente não terá dificuldades para chegar à página que apresenta o jornal recuperado tampouco para

navegar por ela. A palavra-chave (ou a combinação de algumas delas) recuperada no texto aparece sinalizada com uma marcação em verde. Caso haja dúvida, basta clicar no ícone que lembra uma boia e ser direcionado para uma página na “Web” com dicas e explicações sobre toda a navegação do site.

Após a realização da parte prática deste trabalho, focada em necessidades reais de usuários, confirmou-se que conhecer intimamente o mecanismo de busca assim como já ter vivências de navegação pelos resultados é essencial para se tirar melhor proveito da HDB – ou de qualquer uma das hemerotecas estudadas. Por isso, recomenda-se aos usuários que antes de se dedicar ao assunto principal da pesquisa, “passeie” pelas páginas da HDB, experimentando buscas mais livres (procurando notícias sobre sua família, seu bairro, algum fato passado de seu interesse, por exemplo), a fim de se familiarizar com o sistema e suas regras.

Paralelamente a conhecer o mecanismo de busca da informação a ser usado, o pesquisador deve investir um bom tempo avaliando e pensando sobre a sua necessidade de informação. Em outras unidades de informação, em épocas passadas, o consulente contaria com o treinamento, a experiência e a perspicácia de um bibliotecário para percorrer os níveis sugeridos por Taylor, mas atualmente o usuário precisa e deve tomar as próprias decisões.

Desta forma, é recomendável a reflexão sobre o tema da pesquisa (qual é, em que cenário se enquadra?, a quem envolve?), o propósito (elaboração de artigo, dissertação ou tese?, confirmação de um dado?), assim como definir local e período mais propícios à obtenção de resultados. Informar-se sobre o perfil dos jornais TAMbém pode ajudar na seleção do material recuperado. A HDB insere periodicamente artigos sobre os jornais que compõem a coleção, descrevendo o enfoque e o posicionamento.

O usuário deve preparar-se também para tirar o melhor proveito possível dos resultados recuperados, utilizando as buscas já realizadas para melhorar o escopo da necessidade de informação. Como no modelo traçado por Bates, a busca atual levará a outra e outra com melhores resultados até chegar ao que realmente responderá à necessidade de informação – ou até perceber que a busca se esgotou e é hora de redefinir o que se está procurando.

Considerando que nos jornais as notícias tendem a ser construídas a partir de cinco perguntas principais (*O quê? Quem? Quando? Onde? Por quê?*) e que as hemerotecas digitais utilizam em seus mecanismos de buscas pelo menos três

metadados bem definidos (local, título e período), tais informações podem servir para ajudar a definir a *query* que dará início à busca.

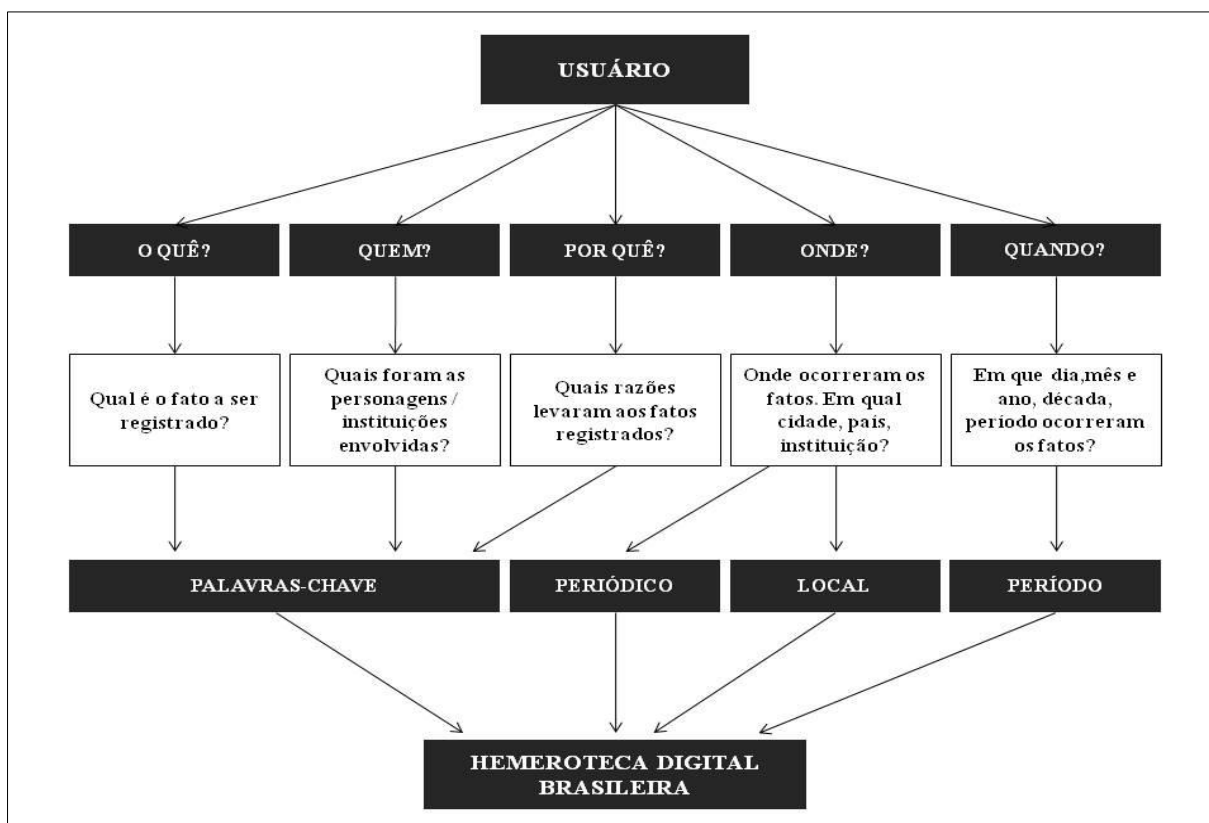
Além disso, as cinco questões podem ser relacionadas às categorias fundamentais de Ranganathan (Quadro 12) para refletir sobre o que pode estar contido em cada uma dessas indagações. O modelo (Figura 69) apresentado a seguir propõem-se a servir como orientação e inspiração aos usuários/pesquisadores que necessitam realizar uma busca em uma hemeroteca digital.

Quadro 12 – Perguntas básicas do Jornalismo alinhadas às categorias de Ranganathan

Categorias fundamentais de Ranganathan	Perguntas principais
Personalidade (<i>personality</i>) – facetas de um assunto que representam sua essência.	O quê? Quem? Por quê?
Matéria (<i>matter</i>) – matérias e substâncias que constituem as coisas	O quê?
Energia (<i>energy</i>) – processos, operações, ações, técnicas, métodos, fenômenos.	O quê? Por quê?
Espaço (<i>space</i>) – divisão geográfica.	Onde?
Tempo (<i>time</i>) – divisão cronológica.	Quando?

Fonte: A autora (2016) a partir de categorias fundamentais de Ranganathan (1963)

Figura 69 – Modelo de conhecimento para reflexão sobre a busca em jornais e revistas



Entre as buscas empreendidas a partir das necessidades de informação dos pesquisadores da área de história da saúde foram extraídos alguns exemplos de aplicação do modelo (Quadro 13).

Exemplo 1: o objeto de estudo do PQ1 é o sociólogo Hiroshi Saito (*quem?*). Por essa razão não haveria como evitar a primeira busca baseada na personagem principal com esta palavra-chave. O período (*quando?*) usado esteve diretamente ligado à trajetória de vida do sociólogo. Apesar de ter vivido em São Paulo, a delimitação do local (*onde?*) não melhorou os resultados. Assim, decidiu-se incluir todos os periódicos de todos os locais disponíveis na HDB.

Exemplo 2: as informações que o PQ2 espera encontrar têm relação com o Serviço Nacional de Doenças Mentais (*quem?*), principal instituição no contexto desta busca. O período (*quando?*) compreende a data de criação do SNDM e as décadas de interesse do PQ2. Por se tratar de um órgão de atuação nacional optou-se por realizar a buscas em todos os estados (*onde?*).

Exemplo 3: a pesquisa é abrangente. Do que foi apresentado pelo PQ4, em uma das buscas, optou-se por usar uma das doenças citadas, “espinhela caída” (*o quê?*) exatamente como era chamada no período da pesquisa, a primeira metade do século XX (*quando?*). Como não havia uma delimitação do local (*onde?*) de ocorrência da doença, escolheu-se buscas em todos os jornais de todos os estados.

Exemplo 4: os protagonistas do movimento verde amarelo (*por quê?*), tema da PQ10, residiam em São Paulo quando o assunto, na década de 1920 (*quando?*) ganhou as páginas dos jornais, mas nas buscas pensou-se em observar se houve repercussão também em outros estados (*onde?*). Assim, entre as buscas empreendidas as que trouxeram mais resultados relevantes foram as que combinaram o nome do movimento com o de um de seus autores (*quem?*).

Quadro 13 – Aplicação do modelo de conhecimento às buscas empreendidas na HDB

	O quê? (Palavras-chave)	Quem? (Palavras-chave)	Por quê? (Palavras-chave)	Onde? (Local / Periódico)	Quando? (Período)
Exemplo 1		Hiroshi Saito		Todos	1919-1999
Exemplo 2		Serviço Nacional de Doenças Mentais		Todos	1940-1979
Exemplo 3	espinhela caída			Todos	1900-1949
Exemplo 4		Menotti del Picchia	Movimento verde-amarelo	Todos	1920-1929

As buscas no acervo da HDB não foram exaustivas, como as que realizaria o pesquisador envolvido na produção de artigo, tese ou dissertação cuja matéria-prima fosse extraída de um jornal. Ainda assim, mais de um processo de busca foi realizado a partir de cada necessidade de informação. O intuito foi observar o funcionamento do sistema, que tipo de informação pode ser recuperada, o quanto o sistema é intuitivo, quais são as possibilidades de filtragem, quais seriam as opções de seguir com a busca a partir do que já foi obtido.

Na maior parte dos casos, em oito das 13, as buscas podem ser consideradas como bem-sucedidas, não apenas porque os resultados aproximaram-se da questão do pesquisador, sendo a busca satisfatória, mas por fornecerem subsídios para buscas futuras e insights para os temas abordados, sendo a busca útil. Realizar buscas demanda tempo, mas mais que isso planejamento, organização e paciência. Os resultados podem suscitar novas curiosidades sobre as épocas em questão, sobre os costumes e as expressões culturais e a formação das cidades e estados. O jornal é um reflexo da sociedade, mostrando comportamentos e desejos de indivíduos diante das notícias que surgem no lugar, o posicionamento político, a maneira de se comportar de determinada época.

Em algumas buscas os processos mostraram-se infrutíferos, com baixa precisão e pouco satisfatórios (como resumido no Quadro 14). Os motivos para tal resultado podem ser desenvolvidos como: inexistência no acervo digitalizado de jornais que cubram o período desejado (PQ3 e PQ6), inexistência de jornais do local desejado (PQ3), informações insuficientes sobre a necessidade de busca (PQ5) ou cobertura insuficiente dos jornais sobre o assunto tratado (PQ11 e PQ13).

Quadro 14 – Resumo das buscas empreendidas na HDB

Pesquisador	Local	Período	Periódicos	Palavra-chave	Precisão (alta, média, baixa)	Resultados (total de ocorrências)	Avaliação dos Resultados (satisfatória, útil, pouco satisfatória)
1	Todos	1920-1999	Todos	“Hiroshi Saito”	alta	143	satisfatória
2	Todos	1940-1979	Todos	“Serviço Nacional de Doenças Mentais”	média	2977	útil
	Todos	1960-1979	Todos	“Divisão Nacional de Saúde Mental”	baixa	333	pouco satisfatória
	Todos	1970-1979	Todos	“nova política de saúde mental”	média	20	útil
3	Pará	Todos	Todos	“epidemia + varíola + Macapá”	baixa	121	pouco satisfatória
4	Todos	1900-1969	Todos	“plantas medicinais”	média	6121	útil
	Todos	1900-1969	Todos	“espinhela caída”	alta	310	satisfatória
5	Minas Gerais	1820-1829	Todos	“cirurgiões + câmara”	baixa	9	pouco satisfatória
6	Todos	1810-1869	Todos	“sangradores + pretos”	baixa	19	pouco satisfatória
7	Rio de Janeiro	1820-1869	Todos	“parteira”	alta	2643	satisfatória
8	Todos	1800-1859	Todos	“erário mineral”	alta	35	útil
9	Todos	1930-1949	Todos	“Instituto Behring”	baixa	30	útil

Pesquisador	Local	Período	Periódicos	Palavra-chave	Precisão	Resultados (total de ocorrências)	Avaliação dos Resultados
10	Todos	1920-1929	Todos	“grupo verde-amarelo”	alta	26	satisfatória
	Todos	1920-1929	Todos	“verde + amarelo + Cassiano Ricardo”		64	
	Todos	1920-1929	Todos	“Cassiano Ricardo”		841	
	Todos	1920-1929	Todos	“verde + amarelo + menotti del picchia”		28	
	Todos	1920-1929	Todos	“menotti del picchia”		1794	
	Todos	1920-1929	Todos	“verde + amarelo + Plínio Salgado”		28	
	Todos	1920-1929	Todos	“Plinio Salgado”		1257	
11	Minas Gerais	Todos	Todos	“Hospital São Francisco”	baixa	71	pouco satisfatória
12	Rio de Janeiro	1910-1949	Todos	“psychanalyse”	média	1618	útil
	Rio de Janeiro	1900-1949	Todos	“psicanálise”	baixa	6287	útil
13	Mato Grosso	Todos	Todos	“alienados”	baixa	100	pouco satisfatória

10 CONCLUSÃO E PERSPECTIVAS DE ESTUDOS FUTUROS

A relação estabelecida com o acervo ao longo de meses, as buscas para este estudo resultaram nas seguintes considerações finais sobre a Hemeroteca Digital Brasileira.

A coleção disponível *on-line* compreende desde os primórdios da imprensa brasileira, em 1808, até jornais contemporâneos que cederam os direitos de leitura à Biblioteca Nacional, como o *Jornal do Brasil*, do Rio de Janeiro, e *O Fluminense*, de Niterói. Historiadores, jornalistas, escritores e outros pesquisadores que busquem informações para retratar qualquer período da História do Brasil pós-vinda da família real certamente encontrarão matéria-prima nas páginas de jornais e revistas digitalizados pela Biblioteca Nacional do Brasil.

Tais informações, antes acessíveis somente no endereço físico da BN, agora podem ser consultadas em qualquer horário do dia, sem interrupções e sem acarretar custos de deslocamento ou mesmo de gasto de tempo necessário para a realização de uma busca de informação.

Manutenção dos mecanismos já instalados e plano de melhorias fazem parte da rotina de qualquer unidade de informação. Nesse sentido, repensar a exposição dos resultados e mesmo a oferta de filtros para refiná-los, a exemplo de outras hemerotecas apresentadas neste trabalho, poderia ser uma maneira de tornar o uso da HDB ainda mais atraente.

Outra sugestão seria o aproveitamento de todas as possibilidades oferecidas por determinado recurso, como o OCR, que entrega aos usuários mais do que um PDF, mas uma versão em texto que pode ser copiada e colada, por exemplo. Este seria um expediente muito útil especialmente aos pesquisadores científicos que precisam copiar o conteúdo de reportagens para utilizá-lo em seus artigos, teses e dissertações.

A HDB apresenta um nível de desenvolvimento semelhante ao das hemerotecas digitais de Bibliotecas Nacionais de outros países estudadas neste trabalho, permitindo o acesso de diferentes tipos de usuários a mais de 15 milhões de páginas de aproximadamente 5 milhões de títulos publicados no Brasil nos últimos 200 anos. Além disso, considera-se que o conteúdo dos jornais e revistas digitalizados pela Biblioteca Nacional apresenta-se a pesquisadores, não apenas das áreas da História ou do Jornalismo, como uma importante e relevante fonte de informação.

As informações coletadas e apresentadas neste estudo podem servir de inspiração para novas melhorias ainda não desenvolvidas.

Além dessas observações específicas para a HDB, a partir da realização desta pesquisa, detalhadas na seção anterior, apresentam-se sugestões para desdobramentos futuros:

- 1) *estudos de usuários de hemerotecas digitais*, a fim de buscar indicações sobre como melhorar interfaces entre os sistemas e os usuários potenciais. Este tipo de trabalho poderia ser desenvolvido não apenas pelos funcionários das unidades de informação, mas por estudantes das áreas de Biblioteconomia e Ciência da Informação, assim como de Design e Informática com foco em Ergonomia e Interação Humano-Computador. Os resultados ajudariam a aperfeiçoar interface e leiaute, assim como os recursos de busca e filtragem dos conteúdos.
- 2) *ferramentas para a representação da informação de jornais e revistas*, encontrando um equilíbrio entre o custoso trabalho manual e o mecanizado. Este talvez seja o maior desafio aos profissionais da informação, pensar em como aperfeiçoar a representação da informação digital, tomando como base os recursos já existentes e usados com a informação analógica, mas procurando elaborar algo que consiga enfrentar os desafios do excesso de informação e dos novos formatos.
- 3) *pesquisas mais detalhadas sobre as hemerotecas* já existentes e a tentativa de desenvolvimento de um sistema reunindo o que há de melhor em cada uma das opções já existentes. Talvez o modelo não sirva para todas as coleções, mas tendo os jornais editados nas últimas décadas formatos similares, talvez um protótipo da hemeroteca ideal sirva para muitas unidades de informação – afinal hemerotecas existem não apenas nas Bibliotecas Nacionais.
- 4) *criação de métodos, treinamentos, cursos e modelos de capacitação* capazes de melhor preparar os usuários para a atividade – agora solitária – de formulação de uma necessidade de informação e da busca na “Web”, mais especificamente a busca em sistemas como os das hemerotecas digitais. Este tipo de treinamento poderia ser elaborado por profissionais que já lidam com usuários em suas práticas profissionais cotidianas, como

os bibliotecários, tendo como base os princípios da competência em informação.

Estudos sobre hemerotecas digitais ainda são escassos e incipientes, assim como os debates sobre como melhor “entregar” este tipo de acervo aos usuários. As experiências estudadas revelam o início de um longo processo. Por meio do retrato dessas 13 hemerotecas digitais espera-se ter contribuído para aumentar a visibilidade desses espaços responsáveis tanto pela preservação da memória quanto pela divulgação e propagação do patrimônio bibliográfico de uma nação.

REFERÊNCIAS

- ALMEIDA, Mauricio B.; BAX, Marcello P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. *Ci. Inf.*, Brasília, v. 32, n. 3, p. 7-20, set./dez. 2003.
- ALONSO NAVARRO, Serafín. La hemeroteca, fuente de investigación histórica. In: MARÍN, Javier et al. (Coord.). *Hemerotecas: aportaciones al estudio y tratamiento de publicaciones periódicas*. Murcia, Espanha: Editora Regional de Murcia/Confederación de Asociaciones de Archiveros, Bibliotecarios, Museólogos y Documentalistas, ANABAD, 1996. p. 36-62.
- ALVARENGA, Lidia. Representação do conhecimento na perspectiva da ciência da informação em tempo e espaço digitais. *Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf.*, Florianópolis, n. 15, 1º sem. 2003.
- AMORIM, Iara Rodrigues. *Competência em informação baseada em inteligência*. 2013. 142 f. Orientador: Jorge Biolchini. Dissertação (Mestrado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Instituto Brasileiro de Informação em Ciência e Tecnologia, Escola de Comunicação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2013.
- APARÍCIO FERNANDEZ, Maria Carmen. Las hemerotecas universitarias, la colección de humanidades y ciencias sociales de la biblioteca de la Universidad de Murcia. In: MARÍN, Javier et al. (Coord.). *Hemerotecas: aportaciones al estudio y tratamiento de publicaciones periódicas*. Murcia, Espanha: Editora Regional de Murcia/Confederación de Asociaciones de Archiveros, Bibliotecarios, Museólogos y Documentalistas, ANABAD, 1996. p. 93-110.
- ASSIS, Machado de. *O jornal e o livro*. São Paulo: Companhia das Letras, 2011.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR12676: métodos para análise de documentos: determinação de seus assuntos e seleção de termos de indexação*. Rio de Janeiro, 1992. 4 p.
- AULA, Anne; NORDHAUSEN, Klaus. Modeling successful performance in Web searching. *Journal of the American Society of Information Science*, Maryland, v. 57, n. 12, p. 1678-1693, 2006.
- BATES, Marcia. Information search tactics. *Journal of the American Society for Information Science*, Maryland, v. 30, n. 4, p. 205-214, 1979.
- _____. The design of browsing and berrypicking techniques for the online search interface. *Online Review*, v. 13, n. 5, p. 407-424, 1989.
- BETTENCOURT, Angela Monteiro. *A representação da informação na Biblioteca Nacional: do documento tradicional ao digital*. Rio de Janeiro: Biblioteca Nacional, 2014.
- _____. *A representação da informação na Biblioteca Nacional do Brasil: do documento tradicional ao digital*. 2011. 154 f. Orientadora: Rosali Fernandez de Souza. Dissertação (Mestrado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Faculdade de Administração e Ciências Contábeis,

Universidade Federal do Rio de Janeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia, PPGCI, Rio de Janeiro, 2011.

BETTENCOURT, Angela Monteiro; PINTO, Monica Rizzo Soares. A hemeroteca digital brasileira. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO, 25., 2013. *Anais...* Florianópolis, SC: Febab, jul. 2013.

BETTIOL, Maria Regina Barcelos. *A Fundação da Biblioteca Nacional: uma memória compartilhada entre dois mundos*. Porto Alegre: Instituto Histórico e Geográfico do Rio Grande do Sul, 2008. Disponível em: http://www.ihgrgs.org.br/Contribuicoes/Fundacao_da_Biblioteca_Nacional.htm. Acesso em: 14 jun. 2014.

BIOLCHINI, Jorge Calmon de Almeida. Semântica e cognição em bases de conhecimento: do vocabulário controlado à ontologia. *DataGramaZero*, Rio de Janeiro, v. 2, n.5, out. 2001. Não paginado.

BOCCATO, Vera Regina Casari. A linguagem documentária vista pelo conteúdo, forma e uso na perspectiva de catalogadores e usuários. In: FUJITA, Mariângela Spotti Lopes (Org.). *A indexação de livros: a percepção de catalogadores e usuários de bibliotecas universitárias*. Um estudo de observação do contexto sociocognitivo com protocolos verbais. São Paulo: Cultura Acadêmica, 2009. cap. 6, p. 119-136.

BORKO, H. Information Science: what is it? *American Documentation*, v. 19, n.1, p. 3-5, jan. 1968.

BRANSKI, Regina Meyer. Recuperação da informação na “Web”. *Perspect. Ciênc. Inf.*, Belo Horizonte, v. 9, n. 1, p. 70-87, jan./jun. 2004.

BRÄSCHER, Marisa; CAFÉ, Lígia. Organização da informação ou organização do conhecimento? In: LARA, Marilda L.G.; SMIT, Johanna. (Org.). *Temas de pesquisa em Ciência da Informação no Brasil*. São Paulo: Escola de Comunicação e Artes/USP, 2010.

BRASIL. Associação Nacional de Jornais (ANJ). *Imprensa brasileira: dois séculos de história*. Brasília, 20---. Disponível em: <<http://www.anj.org.br/imprensa-brasileira-dois-seculos-de-historia-2/>>. Acesso em: 4 set. 2015.

BRASIL. Biblioteca Nacional. *Anais da Biblioteca Nacional*. Rio de Janeiro, 20---. Disponível em: <<http://bndigital.bn.br/acervo-digital/anais-biblioteca-nacional/402630>>. Acesso em 12 jan. 2016.

BRASIL. Biblioteca Nacional. *Apresentação*. Rio de Janeiro, 20---. Disponível em: <<https://www.bn.br/sobre-bn/apresentação>>. Acesso em: 12 jan. 2016.

BRASIL. Biblioteca Nacional. *Depósito Legal*. Rio de Janeiro, 20---. Disponível em: <<https://www.bn.br/sobre-bn/apresentação>>. Acesso em: 12 jan. 2016.

BRASIL. Biblioteca Nacional. *Plano Nacional de Microfilmagem de Periódicos Brasileiros*. Rio de Janeiro, 20---. Disponível em: <<http://www.bn.br/explore/planos-preservacao/plano-nacional-microfilmagem-periodicos-brasileiros>>. Acesso em 12 jan. 2016.

BRASIL. Biblioteca Nacional. *Sobre a BNDigital*. Rio de Janeiro, 20---. Disponível em: <<http://bndigital.bn.br/sobre-a-bndigital/>>. Acesso em: 12 jan. 2016.

BRASIL. Conselho Nacional de Arquivos. Rio de Janeiro, 2010. *Recomendações para digitalização de documentos arquivísticos permanentes*. Disponível em: <http://www.conarq.arquivonacional.gov.br/media/publicacoes/digitalizacao/conarq_digitalizacao_recom_2010.pdf>. Acesso em: 30 ago. 2014.

BROOKES, B. C. The foundations of Information Science. Part I. Philosophical Aspects. *Journal of the American Society of Information Science*, Maryland, v. 3, p. 3-12, 1980.

BÜCHEL, Meinrad. *Sobre a hemeroteca da Biblioteca Nacional de Liechtenstein*. [mensagem pessoal]. Mensagem recebida por <rafaelagiordano@gmail.com> em 15 set. 2015.

BURKE, Peter. *Uma história social do conhecimento I: de Gutenberg a Diderot*. Rio de Janeiro: Zahar, 2003.

BUSH, Vannevar. As we may think. *Atlantic Monthly*, Washington, DC, v. 176, n. 1, p. 101-108, July 1945.

CALONGA, Mauricio Dantielly. O jornal e suas representações: objeto ou fonte da história? *Comunicação & Mercado/UNIGRAN*, Dourados, MS, v. 1, n. 2, edição especial, p. 79-87, nov. 2012.

CAMPBELL, James W.P. *A biblioteca: uma história mundial*. Tradução de Thais Rocha. São Paulo: Edições SESC São Paulo, 2015.

CAMPELLO, Bernadete Santos. *Letramento informacional no Brasil: práticas educativas de bibliotecários em escolas de ensino básico*. 2009. 208 f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2009.

CAMPOS, Maria Luiza de Almeida. *Linguagem documentária: teorias que fundamentam sua elaboração*. Niterói, RJ: EdUFF, 2001. Disponível em: <www.uff.br/ppgci/editais/linguagem.pdf>. Acesso em: 28 fev. 2015.

_____. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. *Ci. Inf.*, Brasília, v. 33, n. 1, p. 22-32, jan./abr. 2004.

_____. O papel das definições na pesquisa em ontologia. *Perspect. Ciênc. Inf.*, Belo Horizonte, v. 15, n. 1, p. 220-238, jan./abr. 2010. Disponível em: <<http://www.scielo.br/pdf/pci/v15n1/13.pdf>>. Acesso em: 27 jan. 2016.

CAPURRO, R; HJØRLAND, B. O conceito de informação. *Perspect. Ciênc. Inf.*, Belo Horizonte, v. 12, n. 1, 2007. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/54>>. Acesso em: 21 fev. 2015.

CARDOSO FILHO, Jair Cunha; SANTOS, Marcia Mazo. Principais aplicações na ciência da informação. In: ALVARES, Lillian (Org.). *Organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações*. São Paulo: B4 Ed.,

2012. cap. 4, p. 185-223.

CARLAN, Eliana. *Sistemas de organização do conhecimento: uma reflexão no contexto da ciência da informação*. 2010. 195 f., il. Dissertação (Mestrado em Ciência da Informação) – Universidade de Brasília, Brasília, 2010.

CARLAN, Eliana; MEDEIROS, Marisa Bräscher Basílio. RICI: R.Ibero-amer. *Ci. Inf.*, Brasília, v. 4, n. 2, p. 53-73, ago./dez.2011.

CAVALCANTE, Maria Juraci Maia. O Jornal como fonte privilegiada de pesquisa histórica no campo educacional. In: *CONGRESSO BRASILEIRO DE HISTÓRIA DA EDUCAÇÃO*, 2., Natal, 2002.

CESARINO, Maria Augusta da Nóbrega; PINTO, Maria Cristina Mello Ferreira. Cabeçalho de assunto como linguagem de indexação. *R. Esc. Bibliotecon.* UFMG, Belo Horizonte, v.7, n.2, p. 268-288, set.1978.

COHEN, Ilka Stern. Diversificação e segmentação dos impressos. In: MARTINS, Ana Luiza; LUCA, Tania Regina de (Org.). *História da imprensa no Brasil*. São Paulo: Contexto, 2008. p. 103-130.

COLE, Charles. A theory of information need for information retrieval that connects information to knowledge. *Journal of the American Society of Information Science*, Maryland, v. 62, n.7, p. 1216-1231, 2011.

COLEPÍCOLO, Eliane; HOLANDA, Adriano de Jesus; RUIZ, Evandro Eduardo Seron; WAINER, Jacques; PISA, Ivan Torres. MeSH: de cabeçalho de assunto a tesouro CONGRESSO DA SOCIEDADE BRASILEIRA DE INFORMÁTICA EM SAÚDE. *Anais...* Santos: Sbis, 2006. Disponível em: <<http://www.sbis.org.br/cbis/arquivos/994.pdf>>. Acesso em: 5 set. 2015.

CORRÊA, Thomaz Souto. A era das revistas de consumo. In: MARTINS, Ana Luiza; LUCA, Tania Regina de (Org.). *História da imprensa no Brasil*. São Paulo: Contexto, 2008. p. 207-232.

CUNHA, Murilo Bastos da. Das bibliotecas convencionais às digitais: diferenças e convergências. *Perspect. Ciênc. Inf.*, Belo Horizonte, v. 13, n. 1, p. 2-17, jan./abr. 2008. Disponível em: <http://www.scielo.br/scielo.php?pid=S1413-99362008000100002&script=sci_arttext>. Acesso em: 8 ago. 2015.

CUNHA, Murilo Bastos da; CAVALCANTI, Cordélia Robalinho de Oliveira. *Dicionário de Biblioteconomia e Arquivologia*. Brasília: Briquet de Lemos/Livros, 2008.

CURIEL DEFOSSÉ, Guadalupe; JIMÉNEZ RIVERA, Ricardo Javier. *Diez años de la Hemeroteca Nacional Digital de México (HNDM): breve reseña de una larga gestión de preservación y acceso: 2002-2012*. In: GALINA RUSSELL, Isabel. (Coord.). *Textos, pixelsy bits: reflexiones sobre la publicación digital*. Ciudad de México: Universidad Nacional Autónoma de México, Instituto de Investigaciones Bibliográficas, México, 2015. Disponível em: <<http://www.hndm.unam.mx/files/acerca-de-hndm/diez-anos-de-la-hemeroteca-nacional-digital-de-mexico.pdf>>. Acesso em: 10 jan. 2016.

- CURRÁS, Emília. *Tesaurus, linguagens terminológicas*. Brasília: IBICT, 1995. Disponível em: <<http://livroaberto.ibict.br/handle/1/454>>. Acesso em: 27 fev. 2015.
- DAHLBERG, Ingetraut. Teoria do conceito. *Ci. Inf.*, Rio de Janeiro, v. 7, n. 2, p. 101-107, 1978a. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/1680/1286>>. Acesso em: 7 fev. 2015.
- _____. Fundamentos teórico-conceituais da classificação. *Rev. Bibliotecon.*, Brasília, v. 6, n. 1, jan./jun. 1978b.
- DARNTON, Robert. *A questão dos livros: passado, presente e futuro*. São Paulo: Cia. das Letras, 2010.
- D'HOORE, Marc; WESE, Veronique; HENRY, Maud. Sobre a BelgicaPress da Biblioteca Real da Bélgica. [mensagem pessoal]. Mensagem recebida por <rafaelagiordano@gmail.com> em 7 set. 2015.
- DUDZIAK, E. A. Information literacy: princípios, filosofia e prática. *Ci. Inf.*, Brasília, v. 32, n. 1, maio 2003. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/123>>. Acesso em: 4 set. 2015.
- ECO, Umberto. Informação demais faz mal. *Época*, São Paulo, p. 46-49, 2 jan. 2012.
- ELEUTÉRIO, Maria de Lourdes. Imprensa a serviço do progresso. In: MARTINS, Ana Luiza; LUCA, Tania Regina de (Org.). *História da imprensa no Brasil*. São Paulo: Contexto, 2008. p. 83-102.
- ESTEBAN NAVARRO, Miguel Ángel. El marco disciplinar de los lenguajes documentales: la Organización del Conocimiento y las ciencias sociales. *Scire*, Zaragoza, v. 2, n.1, p. 93-107, en.-jun. 1996.
- FERNEDA, Edberto. *Recuperação de informação: análise sobre a contribuição da Ciência da Computação para a Ciência da Informação*. 2003. 147 f. Orientadora: Johanna Wilhelma Smit. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicação e Artes da Universidade de São Paulo, São Paulo, 2003.
- FIUZA, Marysia Malheiros. A catalogação bibliográfica até o advento das novas tecnologias. *R. Esc. Bibliotecon. UFMG*, Belo Horizonte, v. 16, n. 1, p. 43-53, mar. 1987.
- _____. O ensino da Catalogação de assunto. *R. Esc. Bibliotecon. UFMG*, Belo Horizonte, v. 14, n. 2, p. 257-269, set. 1985.
- FLUSSER, Vilém. *A escrita: há futuro para a escrita*. São Paulo: Annablume, 2010.
- FONSECA, Edson Nery da. *Introdução à Biblioteconomia*. 2. ed. Brasília: Briquet de Lemos/Livros, 2007.
- FRANCELIN, Marivalde Moacir; KOBASHI, Nair Yumiko. Concepções sobre o conceito na organização da informação e do conhecimento. *Ci. Inf.*, Brasília, v. 40, n. 2, p. 207-228, maio/ago. 2011. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/1856>>. Acesso em: 7 fev. 2015.
- FUJITA, Mariângela Spotti Lopes; RUBI, Milena Polsinelli; BOCCATO, Vera Regina

Casari. As diferentes perspectivas teóricas e metodológicas sobre indexação e catalogação de assuntos. In: FUJITA, Mariângela Spotti Lopes (Org.). *A indexação de livros: A percepção de catalogadores e usuários de bibliotecas universitárias. Um estudo de observação do contexto sociocognitivo com protocolos verbais*. São Paulo: Cultura Acadêmica, 2009. cap. 1, p. 19-42.

GARCIA, Luiz. *O Globo: manual de redação e estilo*. 17. ed. rev.. São Paulo: Globo, 1993.

GIORDANO, Rafaela Boeira. *Da necessidade ao conhecimento: recuperação da informação na “Web” em Ciência da Informação*. 2011. 144 f. Orientador: Jorge Calmon de Almeida Biolchini. Dissertação (Mestrado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Faculdade de Administração e Ciências Contábeis, Universidade Federal do Rio de Janeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia, PPGCI, Rio de Janeiro, 2011.

GLEICK, James. *A informação: uma história, uma teoria, uma enxurrada*. São Paulo: Companhia das Letras, 2013.

GONZÁLEZ DE GÓMEZ, Maria Nélide. Metodologia de pesquisa no campo da Ciência da Informação. *DataGramaZero*, Rio de Janeiro, v. 1, n. 6, dez. 2000. Não paginado.

GRUBER, Tom. *What is an ontology?* [S. l. : s. n.], 1996. Disponível em: <<http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>>. Acesso em: 25 jan. 2016.

GUALLAR, Javier; ABADAL, Ernest; CODINA, Lluís. Hemerotecas de prensa digital. Evolución y tendencias. *El Profesional de la Información*, v. 21, n. 6, p. 595-605, nov./dic. 2012. Disponível em: <<http://diposit.ub.edu/dspace/bitstream/2445/54825/1/619462.pdf>>. Acesso: 8 jan. 2016.

GUINCHAT, Claire; MENOUE, Michel. *Introdução geral às ciências e técnicas da informação e documentação*. 2. ed. corr. e aum. Brasília: IBICT, 1994.

HATSCHBACH, Maria Helena de Lima. *Information Literacy: aspectos conceituais e iniciativas em ambiente digital para o estudante de nível superior*. 2002. 109 f. Orientadora: Gilda Olinto. Dissertação (Mestrado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Escola de Comunicação, Universidade Federal do Rio de Janeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia, PPGCI, Rio de Janeiro, 2002.

HERRERO PASCUAL, Cristina. Tratamiento de publicaciones periódicas. In: MARÍN, Javier et al. (Coord.). *Hemerotecas: aportaciones al estudio y tratamiento de publicaciones periódicas*. Murcia, Espanha: Editora Regional de Murcia/Confederación de Asociaciones de Archiveros, Bibliotecarios, Museólogos y Documentalistas, ANABAD, 1996. p. 63-75.

HOLLEY, Rose. How Good Can It Get? Analysing and Improving OCR Accuracy in Large Scale Historic Newspaper Digitisation Programs. *D-Lib Magazine*, v.15, n. 3/4, Mar./Apr. 2009.

_____. Trove: innovation in access to information in Australia. *Ariadne*, n. 64, July,

2010. Disponível em: <<http://www.ariadne.ac.uk/issue64/holley>>. Acesso em: 18 jul. 2014.

JANSEN, Bernard J.; RIEH, Soo Young. The seventeen theoretical constructs of information searching and information retrieval. *Journal of the American Society For Information Science*, Maryland, v. 61, n. 8, p. 1517-1534 p., 2010.

KEYER ZULOAGA, Graciela. Manual para la organización de una Hemeroteca. *Serie Bibliotecología y Gestión de Información*, Santiago de Chile, n. 44, abr. 2009.

KING, Edmund. Digitisation of newspapers at the British Library. *The Serials Librarian*, v. 49, n. 1/2, p. 165-181, 2005.

LAGE, Nilson. *Ideologia e técnica da notícia*. 3. ed. Florianópolis, SC: Insular, 2001.

LAMPOGLIA, Francis. *Discursividades da/sobre a ditadura militar em uma hemeroteca digital*. 2012. 202 f. Orientador: Valdemir Miotello. Dissertação (Mestrado em Ciência, Tecnologia e Sociedade) – Programa de Pós-Graduação em Ciência, Tecnologia e Sociedade, Centro de Educação e Ciências Humanas, Universidade Federal de São Carlos, São Carlos, SP, 2012.

LANCASTER, Frederick Wilfrid. *Indexação e resumos: teoria e prática*. 2. ed. rev. e atual. Tradução de Antonio Agenor Briquet de Lemos. Brasília, DF: Briquet de Lemos/Livros, 2004.

LANGRIDGE, Derek. *Classificação: abordagem para estudantes de Biblioteconomia*. Tradução de Rosali P. Fernandes. Rio de Janeiro: Interciência, 1977. 1ª Reimpressão, 2006.

LARA, Marilda Lopes Gines de. Linguagem documentária e terminologia. *Transinformação*. Campinas, v. 16, n. 3, p. 231-240, set/dez. 2004.

LAURENZA, Ana Maria de Abreu. Batalhas em letra de forma: Chatô, Wainer e Lacerda. In: MARTINS, Ana Luiza; LUCA, Tania Regina de (Org.). *História da imprensa no Brasil*. São Paulo: Contexto, 2008. p. 178-205.

LE COADIC, Yves François. *A Ciência da Informação*. 2. ed. Brasília: Briquet de Lemos/Livros, 2004.

LESK, M. The seven ages of information retrieval. In: CONFERENCE FOR THE 50TH ANNIVERSARY OF AS WE MAY THINK. Cambridge: MIT, 1995.

LIMA, José Leonardo Oliveira; ALVARES, Lilian. Organização e representação da informação e do conhecimento. In: ALVARES, Lillian (Org.). *Organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações*. São Paulo: B4 Ed., 2012. cap. 1, p. 21-48.

LOPES, Ilza Leite. Estratégia de busca na recuperação da informação: revisão da literatura. *Ci. Inf.*, Brasília, v. 31, n. 2, p. 60-71, maio/ago. 2002.

LUCA, Tania Regina de. A grande imprensa na primeira metade do século XX. In: MARTINS, Ana Luiza; LUCA, Tania Regina de (Org.). *História da imprensa no Brasil*. São Paulo: Contexto, 2008. p. 149-174.

- LUSTOSA, Isabel. *O nascimento da imprensa brasileira*. 2. ed. Rio de Janeiro: Jorge Zahar, 2004.
- MANGUEL, Alberto. *Uma história da leitura*. São Paulo: Cia. das Letras, 1997.
- MANNING, Christopher D.; RAGHAVAN, Prabhakar; SCHÜTZE, Hinrich. *An introduction to information retrieval*. Cambridge, England: Cambridge University Press, April 2009. Disponível em: <<http://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf>>. Acesso em: 4 set. 2015.
- MARTIN, Shannon E.; COPELAND, David A. (Ed.). *The Function of Newspapers in Society: A Global Perspective*. Westport/CT: Praeger Publishers, 2003. Disponível em: <https://books.google.com.br/books?id=ApA8D7XbNqoC&pg=PR9&source=gbs_toc_r&cad=4#v=onepage&q&f=false>. Acesso em: 7 ago. 2015.
- MEY, Eliane Serrão Alves; SILVEIRA, Naira Christofolotti. *Catálogo no plural*. Brasília: Briquet de Lemos/Livros, 2009.
- MIRANDA, Marcos Luiz Cavalcanti de. *Organização e representação do conhecimento: fundamentos teórico-metodológicos na busca e recuperação da informação em ambientes virtuais*. 353f. Orientadora: Rosali Fernandez de Souza. Tese (Doutorado em Ciência da Informação) – Instituto Brasileiro de Informação em Ciência e Tecnologia, Escola de Comunicação, Universidade Federal do Rio de Janeiro, PPGCI, Rio de Janeiro, 2005. Disponível em: <<http://tede-dep.ibict.br/bitstream/tde/54/1/miranda2006.pdf>>. Acesso em: 2 ago. 2015.
- MIRANDA, Silvânia. Como as necessidades de informação podem se relacionar com as competências informacionais. *Ci. Inf.*, Brasília, v. 35, n. 3, p. 99-114, set./dez. 2006.
- MONTE-MÓR, Jannice. Reforma da Biblioteca Nacional. *Ci. Inf.*, Brasília, v. 1, n. 1, p. 15-23, 1972.
- MORAES, Alice Ferry de. A oferta informacional das hemerotecas. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO, 20., 2002, Fortaleza. *Anais...* Fortaleza: FEBAB/ABC, 2002. Disponível em: <<http://www.arca.fiocruz.br/handle/icict/5961>>. Acesso em: 26 jan. 2016.
- MORAIS, Rubem Borba de. *Livros e bibliotecas no Brasil colonial*. 2. ed. Brasília: Briquet de Lemos, 2006.
- MOREL, Marco. Os primeiros passados da palavra escrita. In: MARTINS, Ana Luiza; LUCA, Tania Regina de (Org.). *História da imprensa no Brasil*. São Paulo: Contexto, 2008. p. 23-44.
- MORVILLE, Peter. *Ambient findability: what we find changes who we become*. Sebastopol, Califórnia: O'Reilly, 2005.
- MORVILLE, Peter; ROSENFELD, Louis. *Information architecture for the World Wide "Web"*. 3rd. ed. Sebastopol, CA: O'Reilly Media, 2006.
- MOURA, Dione Oliveira. O relato jornalístico: além do atual, do singular e do extraordinário. In: MOUILLAUD, Maurice; PORTO, Sergio Dayrell (Org.). *O jornal:*

da forma ao sentido. 3. ed. rev. e ampl. Brasília: Ed. UnB, 2012. p. 323-338.

MÜLLER, Christa. *Sobre a hemeroteca da Biblioteca Nacional da Áustria* [mensagem pessoal]. Mensagem recebida por <rafaelagiordano@gmail.com> em 15 set. 2015.

NEVES, Dulce Amélia de Brito. Representação temática da informação e mapas cognitivos. *Inf. & Soc.: Est.*, João Pessoa, PB, v. 22, p. 39-47, 2012.

NISO Z39.19. *Guidelines for the construction, format, and management of monolingual controlled vocabularies*. Bethesda: NISO Press, 2005.

NUNES, Claudio Omar Iahnke. *A análise documentária dos processos de tratamento da informação*. Caxias do Sul, UCS, 10 mar. 2015. 4f. Notas de aula.

PENA, Felipe. *Teoria do jornalismo*. 3. ed. São Paulo: Contexto, 2013.

PEREIRA, Maria de Nazaré Freitas. Prefácio que esclarece o leitor a propósito do sonho de Otlet: aventura em tecnologia da informação e comunicação. In: PEREIRA, Maria de Nazaré Freitas; PINHEIRO, Lena Vânia Ribeiro. *O sonho de Otlet: aventura em tecnologia da informação e comunicação*. Rio de Janeiro: IBICT, 2000, 291p.

_____. Bibliotecas virtuais: realidade, possibilidade ou alvo de sonho. *Ci. Inf.*, Brasília, v. 24, n. 1, p. 101-109, jan./abr. 1995.

PIEIDADE, Maria Antonieta Requião. *Introdução à teoria da classificação*. 2. ed. ver. e aum. Rio de Janeiro: Interciência, 1983.

PINHEIRO, Lena Vania Ribeiro. Informação – esse obscuro objeto da Ciência da Informação. *Morpheus: Rev. Elet. Ciências Humanas – Informação e Sociedade*, Rio de Janeiro: UNIRIO, v. 2, n. 4., 2004.

PORTELLA, Célia Maria. Releitura da Biblioteca Nacional. *Estudos Avançados*, São Paulo, v. 24, n. 69, p. 1-18, 2010. Disponível em: <<http://www.bn.br/portal/arquivos/pdf/celiaMaria.pdf>>. Acesso em: 14 jun. 2014.

QUIROGA, Tiago. Prospecções ontológicas: notícia, acontecimento e o lugar do outro no jornalismo digital. In: MOUILLAUD, Maurice; PORTO, Sergio Dayrell (Org.). *O jornal: da forma ao sentido*. 3. ed. rev. e ampl. Brasília: Ed. UnB, 2012. p. 369-391.

RIBEIRO, Betânia de Oliveira Laterza; SILVA, Elizabeth Farias da; SILVA, Maria Aparecida Alves. Jornal como fonte: uma das pontas do iceberg nas narrativas em história da educação. *Cadernos de História da Educação*, UFU, v. 13, n.1, jan./jun. 2014.

ROBERTSON, S. E. The methodology of information retrieval experiment. In: SPÄRCK Jones, K. (Ed.). *Information retrieval experiments*. London: Butterworths, 1981. p. 9-31.

RODRIGUEZ AMADOR, Laura. *Sobre a hemeroteca da Biblioteca Nacional da Costa Rica*. [mensagem pessoal]. Mensagem recebida por <rafaelagiordano@gmail.com> em 28 set. 2015.

RODRÍGUEZ FUENTES, Lola. Hemeroteca digital de la BNE: tres claves de su éxito. *Métodos de información (MEI)*, II Época, Valencia, v. 2, n. 2, p. 21-27, 2011.

ROWLEY, Jennifer. *A biblioteca eletrônica*. Tradução de Antonio Agenor Briquet de Lemos. 2. ed. Brasília: Briquet de Lemos/Livros, 2002.

RUBI, Milena Polsinelli. Os princípios da política de indexação na análise de assunto para catalogação: especificidade, exaustividade, revocação e precisão na perspectiva dos catalogadores e usuários. In: FUJITA, Mariângela Spotti Lopes (Org.). *A indexação de livros: a percepção de catalogadores e usuários de bibliotecas universitárias*. Um estudo de observação do contexto sociocognitivo com protocolos verbais. São Paulo: Cultura Acadêmica, 2009. cap. 4, p. 81-94.

SANZ, Pascal. La numérisation de la presse à la Bibliothèque nationale de France: sept ans d'action et perspectives de partenariat. In: *IFLA International Newspaper Conference*, Paris, Apr. 2012.

SARACEVIC, T. Ciência da informação: origem, evolução e relações. *Perspect. Ciênc. Inf.*, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996.

SARACEVIC, Tefko. Information Science. *Journal of The American Society for Information Science*, Maryland, v. 50, n. 12, p.1051-1063, 1999.

SARACEVIC, Tefko. In *InCID: R. Ci. Inf. e Doc.*, Ribeirão Preto, v. 1, n. 2, p. 161-167, jul./dez. 2010. Disponível em: <<http://www.revistas.usp.br/incid/article/download/42327/45998>>. Acesso em: 4 set. 2015.

SAYÃO, Luis Fernando. Afinal, o que é biblioteca digital? *Rev. USP*, São Paulo, n. 80, p. 6-17 fev. 2009. Disponível em: <<http://www.revistas.usp.br/revusp/article/download/13709/15527>>. Acesso em: 25 maio 2015.

SCHIESSL, Marcelo; BRÄSCHER, Mariana. Do texto às ontologias: uma perspectiva para a ciência da informação. *Ci. Inf.*, Brasília, v. 40, n. 2, p. 301-311, maio/ago. 2011.

SCHIESSL, Marcelo; SHINTAKU, Milton. Sistemas de organização do conhecimento. In: ALVARES, Lillian (Org). *Organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações*. São Paulo: B4 Ed., 2012. cap. 2, p. 49-118.

SMITH, Alastair G. Internet Search Tactics. *Online Information Review*, Bingley, v. 36, n. 1, p. 7-20, 2012. Disponível em: <www.emeraldinsight.com/1468-4527.htm>. Acesso em: 9 ago. 2015.

SODRÉ, Nelson Werneck. *História da imprensa no Brasil*. 4. ed. atual. Rio de Janeiro: Mauad, 1999.

SOUZA, Renato Rocha. Sistemas de recuperação de informações e mecanismos de busca na "Web": panorama atual e tendências. *Perspect. Ciênc. Inf.*, Belo Horizonte, v. 11, n. 2, ago. 2006. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362006000200002&lng=en&nrm=iso>. Acesso em: 7 mar. 2015.

SOUZA, Rosali Fernandez de. Organização do conhecimento. In: TOUTAIN, Lídia Maria Batista Brandão (Org.). *Para entender a ciência da informação*. Salvador: EDUFBA, 2007. p. 103-123.

_____. Trajetória exemplar. In: BETTENCOURT, Angela Monteiro. *A representação da informação na Biblioteca Nacional: do documento tradicional ao digital*. Rio de Janeiro: Biblioteca Nacional, 2014. p. 11-13.

SPINELLI Jayme; BRANDÃO, Emiliana; FRANÇA, Camila. *Manual técnico de preservação e conservação: documentos extrajudiciais C N J*. Rio de Janeiro: Arquivo Nacional - Biblioteca Nacional, 2011.

STEPHENS, Andy; BRAZIER, Caroline; SPENCE, Phil. A case study in national library innovation: Newspapers in the British Library. *IFLA*, v. 40, n.3, p. 206-212, 2014.

TAYLOR, Robert S. Question-negotiation and information seeking in libraries. *College and Research Libraries*, n. 29, p. 178-194, May 1968.

THOMAS, Deborah. *Sobre o Chronicling America da Library of Congress* [mensagem pessoal]. Mensagem recebida por <rafaelagiordano@gmail.com> em 29 set. 2015.

TRISTÃO, Ana Maria Delazari; FACHIN, Gleisy Regina Bóries; ALARCON, Orestes Estevam. Sistema de classificação facetada e tesouros: instrumentos para organização do conhecimento. *Ci. Inf.*, Brasília, v. 33, n. 2, p. 161-171, maio/ago. 2004.

VALLAS, Philippe. Prepare to digitize: BNF experience. In: *IFLA INTERNATIONAL NEWSPAPER CONFERENCE*, Paris, Apr. 2012.

VICKERY, Brian. *A note on knowledge organization*. 2008. Disponível em: <http://www.iva.dk/bh/lifeboat_ko/CONCEPTS/Vickery_a_note_on_knowledge_organization.htm>. Acesso em: 2 set. 2015.

VIEIRA, Ronaldo. *Introdução à teoria geral da Biblioteconomia*. Rio de Janeiro: Interciência, 2014.

VILLAMÉA, Luiza. Revolução tecnológica e reviravolta política. In: MARTINS, Ana Luiza; LUCA, Tania Regina de (Org.). *História da imprensa no Brasil*. São Paulo: Contexto, 2008. p. 249-267.

WERSIG, Gernot; NEVELLING, Ulrich. The phenomena of interest to Information Science. *The Information Scientist*, v. 9, n. 4, p.127-140, Dec. 1975.

WILSON, Thomas D. Models in information behaviour research. *Journal of Documentation*, London, v. 55, n. 3, p. 249-270, 1999.

YANG, K. Information retrieval on the “Web”. *Annual Review of Information Science and Technology*, v. 39, p. 33-80, 2005. Disponível em: <http://elvis.slis.indiana.edu/kiyang/pubs/”Web”ir_arist.pdf>. Acesso em: 26 jun. 2010.

ZURKOWSKI, P. G. *Information services environment relationships and priorities*. Washington D. C.: National Commission on Libraries, 1974. Disponível em: <<http://www.eric.ed.gov/PDFS/ED100391.pdf>> Acesso em: 13 ago. 2015.